DOCUMENT RESUME

ED 061 299 TM 001 322

AUTHOR	Scriven, Michael; And Others
TITLE	An Evaluation System for Regional Labs and R&D Centers.
INSTITUTION	Ohio State Univ., Columbus. Research Foundation.
SPONS AGENCY	National Center for Educational Research and Development (DHEW/OE), Washington, D.C.
BUREAU NO	BR-1-0857
PUB DATE	31 Aug 71
GRANT	OEG-0-71-4558
NOTE	103p.
EDRS PRICE	MF-$0.65 HC-$6.58
DESCRIPTORS	Advisory Committees; Decision Making; Evaluation Criteria; *Evaluation Methods; Measurement Goals; Program Budgeting; *Regional Laboratories; Specialists; *Systems Approach; Teamwork; *Vertical Organization

ABSTRACT

This report submitted by Advocate Team No. 2 to the U.S. Office of Education, Division of Research and Development Resources (formerly Division of Manpower and Institutions) presents a proposed evaluation system for regional labs and R&D centers consisting of a two-tiered panel organization. The tiers are: (1) A Master Panel--a blue-ribbon advisory group chosen from designated specialist groups; and (2) A series of Specialist Panels that submit site-visit and other reports to the Master Panel. (Author/MS)

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY

ED 061299

An Evaluation System
for Regional Labs and R&D Centers

Advocate team report
Michael Scriven, Chairman
Gene V. Glass
Wells Hively
Robert E. Stake

TM 001 322

A Report Presented to the Division of Research and Development Resources • National Center for Educational Research and Development • U.S. Office of Education • Project No. 1-0857; Grant No. OEG-0-71-4558 • August 31, 1971

Project No. 1-0857
Grant No. OEG-0-71-4558

Advocate Team Report

Michael Scriven, Chairman
Gene Glass
Welis Hively
Robert Stake

Consultants

John Holland
Richard Schutz

August 31, 1971

U. S. Office of Education
National Center for Educational Research and Development
Division of Research and Development Resources

FOREWORD:

Schedule conflicts for some members of the team meant that the meeting of this advocate team came rather close to the total project deadline. Considerations of attention-span, time-pressure and unity of format persuaded the Chairman to use the strategy of a short meeting (2+ days), aimed at brainstorming rather than writing - though with some writing to clarify ideas - followed by a one-man writing marathon. This brought us too close to the deadline to allow a full cycle of corrections and further suggestions by the rest of the team by mail. The Chairman is thus particularly conscious of the need to caution the reader against the supposition that these versions of these ideas are endorsed by everyone on the team.

The team felt extremely fortunate in the services of the project director from the Ohio State Evaluation Center, Diane Reinhard, and the DMI monitor, Ray Rackley, both of whom were very helpful and amiable in circumstances that were often trying (e.g. 10 hour days of meetings and the Biltmore dress rules).

TABLE OF CONTENTS

Chapter I - Introduction 1

Chapter II - Negative Goals 7

Chapter III - Brief Description of the Proposed System 12

Chapter IV - More Detailed Discussion of Proposed System 14

Chapter V - Criteria and Procedures 39

Chapter VI - Budget 56

Appendix A 58

Appendix B 95

I. INTRODUCTION

A first task in preparing an evaluation system is to identify the major options open to the decision-maker being serviced. These options may not have been fully apparent to him, so one advantage of this identification is that his perception of his situation may be rendered more realistic. But this identification is essential for the evaluator since the framework for evaluation is always comparative even if not explicitly so, and the appropriate comparison can only be specified if the decision alternatives are known. Here, the evaluation system is for the benefit of DRDR management, in particular, eventually NIE Management. Major options for DRDR management, obvious enough in this case, include:

1. Cut/Raise total lab and center funds
2. Cut/Raise funds for a particular unit
3. Cut/Raise funds for a particular program
4. Cut/Raise funds for evaluation of labs and centers
5. Cut/Raise funds for evaluation of evaluation program

Associated with each of these options are more specific versions of them, which comprise recommendations or requirements that particular features (e.g., personnel, facilities) be amended, deleted or added.

NIE management will have in addition the 'zeroth' alternative:

0. Cut/Boost DRDR funds

The system described in the following pages is designed to service

these decisions. Within the framework of Dr. Frye's July 1971 conceptualization of the lab-center maturation process, a particular subset of the questions covered by 0-5 above were identified and written into the "Charge to the Advocate Teams"; they are hereafter referred to as the Ten Decisions. However, it became clear the DRDR has got another evaluation task to face besides those involved in servicing the Ten Decisions, namely the so-called "Rite of Passage" evaluation scheduled for early 1972, on the basis of which NIE management will decide, or require DRDR to decide, which programs to take under NIE's wing as the Institute becomes an independent entity. The ROP evaluation is <u>not</u> reducible to any of those involved in the Ten Decisions, for reasons which will be elaborated later. The Santa Barbara team felt that the spirit of its commitment required that it design a system incorporating the Rite of Passage evaluation and it has so done. It should be noted that the Frye framework is in tension with the ROP evaluation, and reconciling the two involved considerable difficulty.

The 'Charge to the Teams' could--on the other hand--be interpreted in a way that would have taken us far beyond an evaluation system. For example, to "specify all steps and procedures leading up to each of the ten decisions" could be taken to require us to specify procedures for <u>generating</u> proposals including informal suggestions. Although we gave some attention to this, we did not feel that anything less than a concerted effort focussed on that problem alone would produce a significant improvement over present procedures. We did feel that such an improvement could and probably should be produced. A more nearly marginal case concerns the RFP approach and its presently constituted alternatives. It is one formalized method for generating formal

proposals. (To understand the difference between it and others, one might imagine offering a prize for the best informal proposal and encouraging entry in that competition from people who have no interest in or facilities for carrying out the proposal.) We have required that the RFP's relevant to our evaluation system be developed with the assistance of the evaluation personnel-system which includes consultant panels as well as in-house staff. We are confident that this first step towards a full reform of the RFP system will yield significant gains. But a full reform proposal--an urgent matter--was felt to lie beyond our immediate task.

Another 'remote' step but related issue concerns the use of output from this evaluation system to generate increased funding. Since available funding determines many of the steps to be taken, this 'solicitation' step could be considered "a step...leading up to...the ten decisions," but we did not go into it. Other management responsibilities including Congressional relations, negotiating, adding 'local focus' requirements for labs, selecting the best 'mix' of cuts/raises given merit rankings, fiscal analysis and supervision, and of course final responsibility. Hence, although the proposed system is more intimately concerned with general policy issues than external personnel systems have been previously, it is by no means duplicating management functions.

The present plan is thus intended as a <u>feasible first stage</u> plan. Further stages are possible as soon as (a) experience with this plan enables improvements in it to be made with some confidence (b) matters such as the two identified in the preceding paragraph are attended to

systematically (developing new procedures for stimulating proposals, and for beginning contract negotiations /~the RFP~/).

An important aim of this plan, besides feasibility, is simplicity. This is a planning proposal--as such, it must be realistic, but if it is over-complex it will not mean the same to all those who must evaluate it and select between it, alternatives to it, and mixes thereof. To be discussed as an option, it needs to be a clear option.

Moreover, there is no good reason for a suggested plan to be monadic. Two suggestions are, ceteris paribus, more useful than one. So we have not tried to produce a spurious unity when we would more usefully mention two or more competing--and to our minds equally feasible--alternatives. But these options are variations on the main theme and they do not obscure its structure; simplicity is retained.

One major analytical step involved identifying the situations where distinguishing between lab evaluations and center evaluations is necessary and the situations where they can be regarded as similar units for evaluation. Another involved classifying the Ten Decisions into categories, the members of which could be considered evaluations of the same type. Considerable simplications resulted from these analyses and from the decision to employ a single master list of criteria for evaluation. These considerations plus a detailed examination of the current activities of the labs and centers also suggest that the present terminology is not only misleading to newcomers but may create unnecessary rigidities in planning. It seems entirely appropriate to

refer to all labs and centers as R&D centers, or ER (for Educational Research) Centers/Units, since (a) the proposed regional focus has almost evaporated--governance apart (b) to the extent an intrinsic difference exists, the labs deserve the title of R&D centers more than the centers (c) the new DRDR terminology is more consistent with this (d) the connotation of "laboratory" is--independently of (c)--quite inappropriate since it refers to an experimental rather than a developmental center (e) any suggestion of crucial differences implied by such different titles is incompatible with the inter-unit variance among either type, which far exceeds any differences between the two types (e.g., NPECE research at CEMREL, materials developed at Wisconsin LRDC) (f) the ambiguous status of new entities like WICHE is better handled by using a single title, general enough to cover it and other variations and (g) the policy centers can either retain that title or accept the moderate slander of the new title.

The cost of an entrenched but perverse terminology is rarely visible to those who have grown up with it, but it may be high; for the 'remoter public' of Congressmen and other agencies may be misled by it just as those who use it may be imperceptibly constrained by it. The transition to NIE is a good time to reform the vocabulary.*

---

*Deeper misapprehensions can be found buried in the terminology of the new ABJ guidelines (July 17, 1971). For example, applied research is there defined as "Research derived from fundamental or basic research" (italics supplied). But if pure and applied research are used as exhaustive categories, this excludes most engineering approaches, an unfortunate position for DRDR since engineering is a paradigm for R&D (e.g., rocket and highway research have scarcely employed basic research for years).

The second step in designing an improved evaluation system, once the basic needs/options have been identified, is to identify the deficiencies of the existing system. To this we now turn.

II. NEGATIVE GOALS

When a system has been devised and instituted by inteiligent and experienced people, and has taken the brunt of some heavy political fire without disaster (albeit with some heat), it deserves serious attention by those who plan to replace it. Now the description of the present plan in, for example, the March 28, 1968, document "Plan for Review and Evaluation of R&D Centers" is hard to fault. Troubles arose in two ways. First, NCERD management was inadequately funded. Second, (and, to at least some degree, consequently) the transformation from the relatively general account of the March, 1968 document to specific procedures was unsatisfactory in a number of respects.

With respect to the first point, this may be as good a time as any to say that the whole R&D program has been heavily burdened by the grossly inadequate support for central staff. It would be worth considering whether cutting out one small center, if the funds could be put into DRDR staffing, would not provide an overall gain for R&D. Should NIE not emerge, such a study should be done by the metaevaluator described in this plan.

With respect to the second source of trouble, the specific implementation procedures 1968-71, it must be faced that any set of specific procedures will generate complaints. The new plan is specifically designed to avoid the 1968-71 difficulties; whether it will generate equally serious ones--perhaps just the equal and opposite ones--should be

determined by a systematic metaevaluation for which arrangements should be initiated the moment a final plan is selected. (Otherwise, the data-collection procedures required will not be set up early enough to avoid disruption and gather what is required in time for an evaluation).

Sources for the following criticisms include interviews with Messrs. Silberman, Frye, Mason and several monitors, the staff of numerous labs and centers and the advocate team. Note; (a) the sequence is not of importance; (b) the present plan is called The Board of Visitors Plan; the term "unit" means "lab or center".

1. The interpanel reliability is highly suspect.

2. The site visits are often too short for adequate data-gathering, e.g., 10 minutes with the Dean to decide on degree of university support for an R&D center: or R&D output not even read through, because of time shortage.

3. Too few Board members are specialists in the program areas of a unit with several programs.

4. Little attempt is made at serious comparison of unit with non-USOE R&D units, yet this is the proper comparison.

5. No effort is made to face the "shuffling" problem. (i.e., the problem of showing that the director of a unit is not just shuffling the available manpower around and dealing himself a good hand, perhaps because of the work conditions he can offer, but <u>not adding to the nation's resources</u> of manpower. In short, would the research for e.g., an R&D center have been done (a) without creating the center, i.e., without overhead and management costs (b) without spending any money at all. (Under (a), we would be

considering the alternative of direct funding to individuals or small groups of individuals instead of 'middlemen' funding). The threat of the "shuffling" hypothesis is less for labs which have actually produced something, since the private sector has done little serious R&D.

6. Overviews of unit or program redundancy, need for amplification, increased DRDR staffing, etc., cannot be credible from NACEL (which spends too little time on them) or from DRDR staff (conflict of interest and undermanned), so the credibility of the whole enterprise suffers.

7. The monitor (program associate, program officer) role is schizophrenic, between helper and judge.

8. The supposed attempts to recognize good unit management by a hands-off policy are constantly subverted by frequent survival reviews.

9. The requirement of "mission integrity" for a unit has often been used punitively, although (a) it is inconsistent with OE practice (CEMREL), (b) it can be too easily met by adding mythology to the program plans.

10. The reports of site visitors are too often (not always) too bland or jargonistic; either fault may come from anxiety about losing friends or face in preliminary presentation of the evaluation at the end of the visit.

11. Site visits are extremely disruptive or (if this is inevitable) they occur too often.

12. (a) The monitors have a tendency to pass on low-reliability gossip about new emphases, criteria, interests, etc., which distorts the operation of a unit. (b) When they are right, i.e., standards are

about to be changed in midstream, the results are equally unfortunate.

13. The selection or disbarring of Visitors by the unit director contaminates the evaluation system.

14. The actual reports often exhibit internal evidence of unreliability. For example, guesses about future policies or political events on which the evaluation is made to depend. Even good educators are not good prophets. For example, complaining that Ph.D.'s from weak institutions are on the staff when adequate data exists on which to judge their actual performance. For example, complaints about "inadequate dissemination" when even the developmental budget is marginal. For example, reversals from one year to the next on democratic procedures in management. For example, out-of-the-blue complaints about community contacts.

15. The visit is not as useful to the unit as it should be i.e., it is poor in its role as formative evaluation, e.g., with respect to decisions--implications of recommendations/commendations.

16. Poor awareness of the lab/center differences with respect to "cushion," i.e., no faculty appointment fall-back position for lab staff--<u>hence</u> time-scale for decisions must be different.

17. Intervisitor reliability low <u>in spite of</u> discussions.

18. Boards are university-dominated despite R&D resources in IBM, Xerox, Battelle, SDC, BRL, SRI, DOD.

19. Political standards have constantly intruded over professional ones, e.g., WICHE, Georgia: or attempts to intrude them have occurred, e.g., ERIE.

20. No serious scientific attempt has been made to learn from the evaluation <u>why</u> the failures fail. But that is throwing away

valuable data in an anti-R&D way. If resource-building is the aim, you probably learn as much from failures as from successes.

III. BRIEF DESCRIPTION OF THE PROPOSED SYSTEM

The recommended evaluation system is a two-tiered panel organization with a single Master Panel on top, to which a series of Specialist Panels report. The MP is a blue-ribbon one-year advising group, carefully chosen from designated specialist groups, achieving consistency by preliminary 'calibrating' exercises, by constant composition across the year's evaluations, by one-third personnel carry-forward, and by the use (and improvement) of standardized guidelines and special 'segregated grading' procedures. It does not perform site visits to evaluate labs or centers (hereafter called 'units'): it uses reports from Specialist Panels who site-visit groups of comparable programs only, reports from 'monitors' (whose role is reconceptualized) and reports from a new in-house full-time Support Team. In turn, it reports directly to (a) DRDR, (b) The Director or Commissioner*.

Note A:

Many elements in the plan are simply articulated modules which could be dropped or adopted independently of the rest. (e.g., the metaevaluation system). Others are functional because of interaction

---

*In this respect, the Master Panel occupies a slot like that of the defunct National Advisory Committee for the Educational Labs. But its task and time requirements are different, being better thought of as those of a Standing Site-Visit Committee.

effects (e.g., the allocation of tasks as between Master Panel and Specialist Panels). An attempt is made to identify the latter as such in the text, to aid the synthesizing project.

Note B:

This plan is experimental in the sense that it is sufficiently different from the previous one to represent an exploration of the possibility space. The experiment will be valuable, even if the plan is not, just to the extent that the plan itself is adequately evaluated in practice. For this reason, the role of the metaevaluation module is to be regarded as an investment not a gimmick.

Note C:

The plan is usable regardless of the level of support from Congress (Brickell's question at the briefing session). It requires no serious modification to handle moderate levels of 'external' funding for some units.

Note D:

In a budget of $34 million, a case could be made for $3.4 million in evaluation costs. This plan means slightly over $.34 million plus metaevaluation at about $.034 million.

IV. MORE DETAILED DISCUSSION OF PROPOSED SYSTEM:

The proposal herein meets all of the objections of Section II, not with a watertight guarantee that they will not occur again, but with a specific arrangement that promises improvement. Elaboration begins by confronting an important possible objection, since if postponed a negative set may distort perception of the system details, and since it provides a useful cameo of one aspect of the system in operation.

The system has one feature which runs counter to some expressed desiderata for a new system. The Master Panel does not (in general) site-visit. The trade-offs are:

1. Each program is site-visited by a Specialist Panel reporting to the MP.

2. A single MP can handle all the evaluation in the year, if it doesn't site-visit, yielding (we may expect) great gains in consistency/ reliability.

3. The Special Panel can more easily be helpful to the unit director since it is not, so to speak, the final judge of the whole unit (nurturant role with respect to unit). And it can stay longer than the unit can tolerate a Board of Visitors since the disruption will typically (not at SWRL, a single-program unit) involve only part of the unit and can be more easily conceived as a consultation (nurturant for program). Certainly there will be no loss of credibility-checking ("Are these reports written by these people?", etc.).

4. Each site (except SWRL) is visited by several teams (Special Panels) each of whom will be making incidental comments on unit-management and direct comments on program performance, which is eventually the best indicator of management skill. And the carefully rigged snow-job for the B of V is less likely to hold up over the cummulatively greater period of Special Panel interaction.

5. Each Special Panel typically visits several programs in its area of special knowledge. Hence the comparative role is better served than at present. And the Special Panels will acquire more consistency from making several visits as a group (as well as from special calibration preparation).

6. Given adequate 'security precaution' site visits somewhere in the system, there is much to be said for keeping the MP off-site; the well-known errors due to personal charm or antagonism are truly serious.

But the plan is modifiable in this respect. The Rite of Passage evaluation in 1972 will, however, place the heaviest demands on time, just when a new system is starting up. For that year, off-site evaluation by the MP seems essential.

Working from the June 3, 1971 document DMI Institutional Support and Evaluation Policy, one might make a case in the abstract for approximately fourteen different kinds of evaluation. On the face of it, evaluation of requests from Phase III units for facility grants, of planning grants from pre-institutional groups and of fee use by a Phase III unit are very different enterprises. It would be an easy step from these to the conclusion that one needs 14 instruments and/or types of panels. The logistics of

co-ordination would be staggering, and the logic of relating their conclusions would be baffling. Given that a system has to be, not just fully staffed, but in full swing within a year, it would require someone remarkably optimistic about practical development to encourage such complexity.

The commitment of the Santa Barbara team was very strongly towards the advantages of a unitary system in the first place, perhaps developing extra arms when it has learnt to use its own. Both for logistical and for logical reasons this was thought to be not just making a virtue of necessity but more virtuous. The logistical simplicity is obvious; but the other point is as strong, for it is hard to see what set of instructions could achieve an intergroup consistency comparable with the interdecision consistency of a single panel. The whole set of decisions and evaluations are so closely interactive that a group making one evaluation would have to know almost everything known to the groups making the others. Hence the Master Panel plan, which begins by having one group make most of the decisions--with the charge that it constantly attempt to formalize its procedures so that explicit comparisons with other groups using (a) these formulations, (b) their own approach, may be instituted within 18 months of the beginning of the new system.

This report is written for the DR^2 Staff, and perhaps for a convergence team; it is not in the form in which a final plan would be sent out to units. In <u>that</u> form, the composition of the MP would be given early, for evaluation is a life-and-death matter for a unit and the names (or at least rank and serial number) of the judges in that trial are of prime interest.

But in this presentation it would be quite wrong to give details of the MP first, for clearly the MP cannot be picked well or badly until its role is clearly understood.

Its role can be filled in first by looking at its involvement in the Ten Decisions list, then by discussing the criteria employed, then the training procedures ("calibration"), and finally by looking at selection suggestions. There are certain differences in clarifying the Ten Decisions due to minor errors and unclarities in the June 3, 1971 document which is supposedly the definitive basis for the task of the advocate teams. Doubtless these have been picked up in the review process, but they account for some discrepancies between the ensuing descriptions and that document. (For example, full-fledged BPP's come not only from mature institutions (p.22) but from Phase II institutions (p.14); the milestone review uses evaluators variously described both as "designated" and merely as "approved" by OE (p.23)--the latter seems intended, but no operational definition is given of the distinction (severe standards for approval constitute designation). In other respects, there are divergences from the June 3, 1971 document of the kind that are natural in any refinement process--but they are fewer and less serious than the team had anticipated.

Evaluations Servicing The Ten Decisions (Decision-Clusters)

Decision Cluster Number	Description	Evaluation Agent
1.	Institutional Planning Grant Proposals	MP + Consultants

Decision Cluster Number	Description	Evaluation Agent
2.	Start-Up Grant Proposals	MP + Consultants
3.	Start-Up Performance Review (Phase I)	MP + Specialist panels Including Management Panels
4.	Early Years Review Review (Phase I)	MP + Specialist Panels Including Management Panel
5.	New-Program Planning Proposal	MP + Consultants
6.	New-Program Implementation Proposal	MP + Consultants
7.	Milestone Check*	Monitor + Technical Colleague + Unit Consultants
8.	Completion Check	MP + Specialist Panels
9.	Intervention	MP + Specialist Panels
10.	Budget Adjustment	MP + Specialist Panels

The Technical Colleague is a cross between a monitor and a consultant and represents a further assistance in improving the monitor's utility to DR^2 and to the project. The concept comes from the Belmont Project where it proved very helpful. It grew from the recognition that for some projects at some stages the monitor was not technically able to advise or

---

*The plan for the milestone check requires special comment. There does not seem to be any justification for disguising the fact that the monitor is performing informal "furlong-stone" reviews. While sympathizing with the attempt to simplify the monitor's role-tension problems, the fact remains that he or she is the best-informed DR^2 person about the project and must report discrepancies. The milestone point is a good one at which to formalize this slightly and take some of the burden of the joint role off his/her shoulders.

to detect short-falls in the program. At this point she or he was empowered to bring in a technical aide, differing from a project consultant in being free (paid for by OE) and in the obligation to report to the monitor as well as the project. The difference from the usual OE Consultant was the longevity of the relationship, enabling the individual to become thoroughly familiar with the project. But co-option scarcely becomes serious, not only because of the source of funds, but because the relationship is still intermittent by comparison with the Technical Colleague's basic employment.

It seems plausible that provision of Technical Colleagues funds and reduction in load for the monitor, plus clarification of the evaluation structure will suffice to transform their tasks and performance. (The reduction in load should be coverage of two units in depth; two others slightly; [in case of illness etc.] and one area, e.g., early childhood.) The selection of Technical Colleagues should be done by the monitor from the updated lists used for Specialist Panel selection, and it should get the pro forma* approval of the program manager, as well as (where different from) the unit director on the one end, and the head of the Lab/Center Branch (June 3, 1971 document p.29, notes 4 and 5) at the other.

The milestone review will thus involve the monitor and a consultant of his/hers as well as unit consultants, pro forma approved by DR^2. It is

---

*The acceptable reason for disapproval would be demonstrable conflict of interest (financial, not intellectual), and the general policy would be that a demonstration of this would have to be submitted in writing to DR^2 for consideration, as a response to the letter requesting approval of a named Technical Colleague (alternates not offered).

not suggested that every monitor will have to employ and train a Technical Colleague for substantial use in every program; but as a milestone check comes up it will be necessary to contact one and key him/her in to the background.

The emphasis on DR^2 representation at milestone checks is contrary to the intention expressed in June 3, 1971. There is another reason for it, besides the recognition of actuality as far as the monitor is concerned. That is a recognition of weakness in the early-warning system for intervention reviews. Granted that these are not going to be common in mature institutions, they are possible, and demonstrated failure at a milestone is a far better basis for intervention than the lone report of a monitor. The essentially self-monitored concept of milestone reviews in June 3, 1971 seems rather too 'hands-off'. Even though there are outside consultants involved, they are selected and paid by the unit and approval by DR^2 is likely to amount to a nod of the head by a monitor, whose status is not high or secure enough to make uneasiness readily felt.

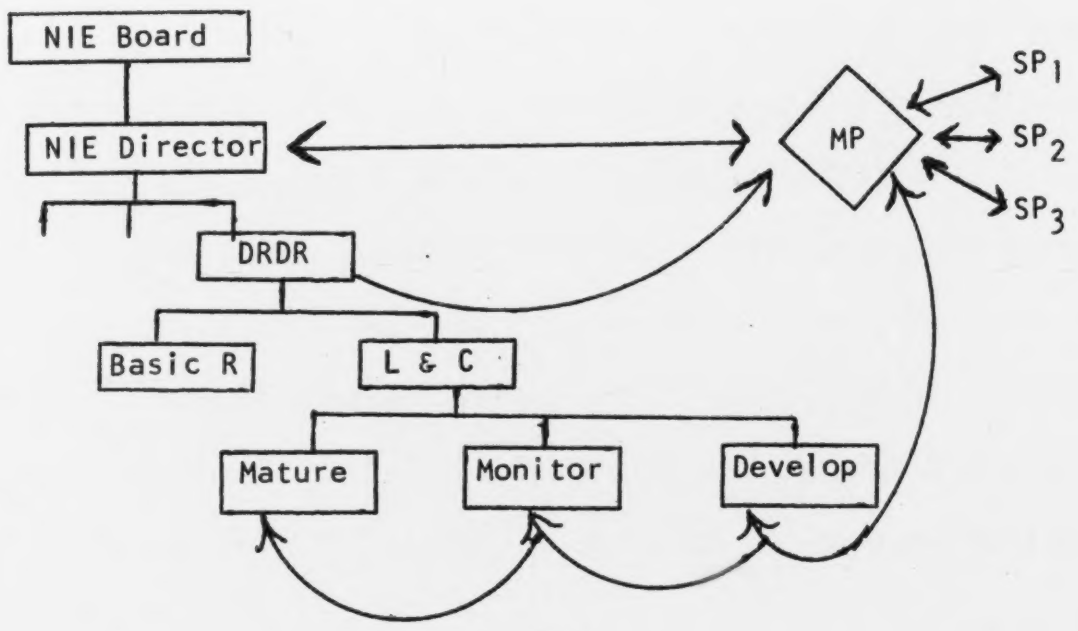
Let us return to the more common kind of review, which typically involves the MP $\pm$ (with or without--at its discretion) a Specialist Panel. The Specialist Panel is a multiple Technical Colleague except that it is an auxiliary of the MP, which collaborates in its selection. Technical Colleagues are an auxiliary of the nurturant arm and may be needed rather quickly-they can be appointed in-house, even though their reports will be seen by the MP (the MP has some indirect control over Technical Colleague selection, since it controls the Specialist Panel list from which Technical Colleagues are drawn).

The most significant weakness of the present system is the lack of fully informed, highly qualified educational professionals at the overview level. In cash terms, the largest possible savings are not available because there isn't enough clout in the system to support a decision to fold half-a-dozen or two dozen programs on the grounds of redundancy with external or internal efforts. It is only because of reduced funding or flagrant failure that cuts have been made. But new natural needs, new perceptions of the texture of the educational discipline, and new personnel distributions may require shifts in emphasis year-by-year. No Advisory Council has enough grasp of the fine structure of the system to justify a really tough line on what needs to be trimmed to make room for the new, or to release funds to double the power of the attacks on the other problems already under way. The MP will combine overall knowledge, acquired during its intensive training and by virtue of its involvement in each kind of evaluation, with extensive external experience and knowledge.

On the other hand, there are important safeguards built into the system against possible idiosyncracies of a new MP. First, there is the very important provision of a one-third carryover in personnel from one year to the next (not so large as to swamp the new group nor so small as to be merely a personal view-from a committee of 10-11, 3 or, rarely, 4 will carry over; at least 6 will be new.) Second, there is the time lag in effectuating decisions which means that in almost every case of a major decision there will be some chance of reconsideration and salvage by the succeeding MP. Third, there is the purely advisory status of the MP, which will allow NIE to temper its recommendations in the light of staff views where conflict emerges--at the risk, of course, of resignation by the MPs.

The balance between arbitrariness and quick reaction-time is a critical parameter for any management system, obviously enough, and it is there that a one-year panel with one-third carryover represents a good balance. Two- to five-year tenure was discussed, as was making the MP a staff panel or mainly a staff panel of DR^2. It should be noticed that all of these point in the same direction by comparison with the present system of low-time-commitment NACEL and narrow-range-of-experience Boards of Visitors (the Chairman typically being involved with three units.) Minimax strategy suggested making the smallest truly significant move in this direction and then evaluating the change--hence the one-year term.

Similar considerations affected the recommendation of a slot in the organization chart. To put the MP below the director of DR^2 leaves the director of NIE with the virtual necessity of establishing an Advisory Committee to translate his board's recommendations into specifics. The MP can certainly incorporate that task into its workload with scarcely any effect. Hence, despite its very direct connection with DR^2 management, we can think of it as serving the Director of NIE as follows:



Consider Next the Other Dimensions of Feasibility of the MP Concept.

Time-Commitment and Recruitability. Although the MP will not consist entirely, perhaps not even mainly of university faculty there is a particular problem about the latter that must be faced immediately, and that is the problem of incentive. Consulting income is normally on top of regular income, bonus; but it is normally limited to the equivalent of one day a week. The MP commitment requires more than that amount of time. But a split appointment--say, a half-time one--offers no renumeration at all, or very little, since the arrangement is normally made on the basis of replacing base salary. In fact, given the traveling involved, it would not be an attractive option. The Advocate Team was divided on whether it would be marginally feasible or definitely not feasible. But we propose a novel plan that improves on it.

The need for initial training is obvious and the nature of it is discussed below. The MP would be convened for a two-month Training Camp in the summer, the one time when the academic can get away for a solid spell without adjusting teaching duties. In fact, only the first week or two of this period would be devoted to training simpliciter i.e., to background reading and briefing and discussion of the situation as it is. (Further details of possible content of the early weeks will be found in the section of the Specialist Panels, who get ultra-compressed versions of the 'basic training' course.) Thereafter, a gradually increasing load of actual evaluation would be phased in. The sequencing of this would be such that relatively

unimportant decisions would be made at first, and even these would be 'sat on' for later consideration by the group. As each new type of evaluation was entered on, the need for new material would emerge and the Support Team or the MP itself would go about getting it or arrange that it be obtained. At this stage, interviews with ex-USOE personnel, directors of terminated units and of non-governmental R & D units would be useful, possibly a site-visit or two,--and intensive briefings by BOB accountants and some unit fiscal officers. Simulations of past and possible decisions should be set up and the results compared with reality where past records were available. And the beginnings of a formalized system would be attempted. The Advocate Team feels that a training program for this period is a crucial need, that it cannot be RFP'd in sufficient time, at a reasonable cost with much expectations of merit given the archival limitations; hence, this first time, it should either be directly contracted for or assumed as the first task of the Support Team. It will be greatly improved if part of the contract includes correction in the light of the first Training Camp.

Hence the MP will be asked not only for drafts of the Operations Manual for MP's but for advice on the Training Manual for MP's. The last weeks of the Camp will consist in rechecking earlier decisions before passing them on and in making full-scale major decisions. Further, the group will settle on details of the operations schedule for the coming nine to ten months, in consultation with the support team.

During that nine-ten month period, the group will revert to what might be called Sustaining Consultant basis. That is, they will undertake to commit 3-4 days per month to this task, essentially eschewing other

consulting work. The normal arrangements for a month are suggested as:

1-2 days monthly meeting at which materials sent out two weeks earlier are discussed, and new materials distributed.

1-1½ days work at home in the first two weeks of the next month, resulting in completed evaluation forms being returned to the support team or a conference call.

½-1 days work in the second fortnight as preparation for the monthly meeting.

This arrangement leaves open the possibility of a month's wind-up and de-briefing session, to round out the year; this slice of the second summer may not be necessary in later years, but is strongly recommended for the first one. At least a two-week period should be scheduled. This would serve the further purpose of providing an excellent opportunity to phase-in the new MP.

Incidentally, every effort should be made to avoid having the Summer Camp in Washington, despite slightly increased logistical problems of alternative locations (the atmosphere will be hot-house enough indoors). If arrangements can be made to go somewhere that MP's families can enjoy as a holiday, the recruitment ration is likely to respond favorably.

It is thought that such an arrangement is workable and attractive for faculty; and no less workable or attractive for non-faculty.

The alternatives were discussed; it was felt that recruiting would suffer heavily if full-time leave was required, especially if moving a family was

involved. On the other hand, mere consulting time would be entirely inadequate. Looking at the work-load suggests that it can be handled even in a Rite-of-Passage year, by the combination plan.

In drawing up a contract with a trainer, as in RFP-ing instruments required elsewhere in this plan, it is suggested that the basis for agreement be a detailed understanding of the plan as a whole, resulting in a proposal to be judged on its utility for the plan as a whole. It is not as satisfactory to attempt in vacuo formulations of the requirements. The RFP or the contract should be of the form "Design a training program for the first Summer Camp of the MP plan, implement it, and correct it in the light of the first summer."

If the time commitment above is enough for the job and not so much as to exclude good recruits, are there other features of MP-membership that should be clarified and that bear on recruiting? The position has prestige and this should be of some value in negotiation with the parent institution of the MP, if any. It might be good practice, in the letter of invitation, to explain that-unless the invitee protests- his acceptance will be followed up with a letter to his chairman, copy to his dean and president, officially requesting his help and explaining the job and its importance. The position involves making a significant contribution to the Nation's education for reasons well expressed in various documents provided to the advocate teams, and elsewhere... There will be a remarkable opportunity to learn about educational R & D... For many members, the training in system-evaluation will be of great value not only to themselves, but to the parent institution (a point which should also be stressed in the covering letter)... The materialistic side for the indiviual will doubtless often appeal to his

interest in powers as well as paying off the mortages; for the institution (which in some cases have to approve exceptions to teaching and/or consulting requirements) there is the chance of acquiring an informant well-versed in the operations of the new NIE or DRDR. For institutions other than universities, very similar considerations apply.

The Chairman of the MP. There is one exception to the above generalities about the MP's role. The Team felt that a kingpin role should be played by one man on the MP, the chairman. His selection should be made with great care, with direct involvement of the Director of NIE (or Commissioner), his appointment should be full-time and his expected time in Washington either full-time or considerably more than that of the other panelists. Once he is selected, he can assist in the selection of the other panelists, including the option of a half-time assistant chairman if he strongly so desires. The thoughts behind this suggestion included the possible advantages of the chairman's attendance at Congressional Committee hearings, as witness or listener, (later in the year he might brief Congressional aides on the new system) his utility for other divisions of OE/NIE, his power in keeping laggard panelists up to schedule, and above all, the elasticity against poorly predictable time-demands on the MP of a key full-timer. In later years, this may well prove quite unnecessary, but a disaster might be avoided in the first year by the extra cost involved. Whether he elects to move to D.C. or stay at home, he should be provided with a secretary.

An option in the selection sequence, this first year, would be to appoint the Chairman before appointing the in-house extra staff that will be required, so that he could assist in those appointments. Marginally, he might assist

in interviewing or dossier-sorting for the new monitors whose appointment is a necessary condition for any plan. Certainly he should play an important part in arranging for the Training Camp. It is not essential that he have had a great deal of experience with the labs and centers, and it is not enough that he be a prestigious figure. The critical property is commitment to developing a refined system of evaluation, one that should be a model for every agency and large company, and the intellectual equipment to go with it. In short, Ralph Tyler or Ralph Tyler the Younger, not a spavined elder statesman who will perform benignly and perhaps well, but leave behind him no recipe for the quantum jump to a new level of efficiency that must take place in Year Two.

The Specialist Panel Role. This is implicit in the above and the task is one which has been quite well performed by individual consultants and committees for some time. The main differences from common practice are: the stability of the panel across several program-evaluations, usually involving site-visits; the demand for formalization of procedure and other steps towards improvement of the process; and some pre-training. Refining the latter cannot be accorded as high a priority in Year One as the matters already discussed, but it should become one of the most important and distinctive features of the evaluation system. A primary version of the calibration training for SP's would involve elements that would form an early unit(s) in the MP Training, viz.,(in this order):

A series of case-studies from NCERD files, supplemented very cautiously with hypothetical cases, selected to exhibit most clearly the characteristics of:

(a) programs that failed, for an assortment of reasons covering the most important 'traps' for program management.

(b) programs that were clearly acceptable though suffering from significant defects.

(c) programs that fooled the evaluators, either because of misleading rhetoric, incorrect statements or subtle inconsistencies. Some of these could have been turned up only by a site-visit, and the panelist must begin to develop a sense of what must or can and cannot be picked up on site-visits.

(d) special test programs. Note that under (a), (b), and (c) there would first be a series of analyzed examples and then a series without analysis which the prospective panelist would then try to analyze. Under (d) he would receive a random sample of programs, some evaluated, some not, and he would be expected to regard each with suspicion, though he would be informed that some were examples that appeared to his predecessor to be paradigms. Eventually it is to be hoped that the DR^2 pool for MP-SP work would be enlarged by asking possible members to take the course and test battery. In addition, DR^2 should work towards an audio-visual repertoire of video tape recorded site visits, (possibly some staged) to develop sensitivity to crucial cues and immunity to (empty) charms.

Some mini-courses in management principles and evaluation methodology would be well worth trying. Content analysis of transcripts of the discussion by earlier site-visit panels should be undertaken by the contractor developing this training program, to identify most-asked questions and most-time-consuming background. More important would be the same evaluation of the sessions following the Year One training.

An extremely serious methodological problem arises in connection with

the training of MPs and SPs though the SPs may not notice it. This is the 'peaches and pears' problem--the problem of comparing unlike entities. On the one hand, it is quite anxiety-provoking to ask scientifically-trained people to do it and they protest that there is no rational way to do it; on the other hand, if they are to administer funds in almost any way, they will have to do it. Even the most isolated academic eventually has to choose one graduate student over another for a job or a fellowship when their skills and personalities lie along entirely different dimensions. At first sight, the SPs seem designed to avoid the peaches and pears problem. They compare similar programs, and certainly they will be unlikely to protest the impossibility of their task. But in reality, they will not be looking at different groups taking the same exams. They will be looking at very different approaches to very different problems with very different resources and several dimensions of success on all of which the groups are widely spread.

It can in fact be argued that evaluation always involves the peaches and pears problem. Suppose your task is to compare the performance of several runners, on the same day on the same track in the same conditions with respect to speed alone. The required experiment is called a race, and timing it accurately is a minor exercise in applied science, one which would never be referred to as evaluation. Evaluation necessarily involves the holistic rating of disparate elements. The way we are trained to do it rightly encourages us to reduce the variance as much as we can. But it wrongly suggests there can be no reliability about doing it across widely disparate entities. As S.S. Stevens discovered when his colleagues said subjects couldn't reliably quantify loudness etc., logic doesn't limit us

as much as scientists think. The experience of good administrators refutes much of the peaches and pears pessimism; but its existence is something a training program must reduce because it effectively inhibits the development of any skill. The major difference in the task of MPs and SPs, aside from time-requirements, is the level of generality at which the comparisons must be made. The MP training program will have to focus on this part from the beginning; the SP program can, almost, ignore it.

The Support Team Role. Implicit in the preceding is an account of much of the ST's work, but we can now elaborate briefly. Apart from the major tasks of developing the lists of resource personnel for both SPs and MPs (and hence Technical Colleagues), collecting and processing input material for the panels and consultants, handling contracts for instruments and training, making their own reports on how the evaluation process is proceeding and suggesting alternatives to it, reporting out to DR^2 and NIE management, constructing and maintaining the very tight schedule for the post-camp period, the ST will undergo the same training program as MP and SPs and take part in the intensive efforts to improve it.

The ST should have no other duties within DR^2, certainly not in Year One. It should have its own quarters and the Advocate Team liked Stake's idea of a 'War Room' where massive displays of the present and past status of every program and unit would be kept up-to-date, where ST staff would be available for inquires from within and without NIE as to present status, products and research, etc., and where briefing/debriefing can be done. There should be as complete a library as possible in this or an adjacent room, with copies of all important documents from the units, as well as staff documents about them. Someone on the ST should be designated as Display Officer.

someone with strong feeling for the mighty importance of data-presentation format/media etc., and with influence extending to editorial powers over DR[2] publications concerning the units. Someone else should be designated as Archivist, with an equal sense of the extent to which cummulative science depends on its record-keeping. It would be preferable if these individuals could be appointed with these duties in mind. The total size of this staff need not be great. Our suggestion was one senior Ph.D. with evaluation experience, about Grade 15; two M.A.'s in the social sciences (roughly); and one or two interns or junior aides; plus a secretary and a part-time librarian, assuming some access to a secretarial pool for high-concentration periods.

The Monitors. They would not be part of the ST, partly because of the desire to de-emphasize their evaluation role and encourage their advocate role. The 'advocate role' should be defined, not as emotional liking for 'their' programs, but as the capacity to make as good a case as possible for them; if the monitor can also make an excellent, perhaps overwhelming case against them, that is no disqualification (if the case is sound), any more than obvious emotional identification with them. The crucial requirements are understanding the program which is a prerequisite for being able to represent it in Washington. This identification of the requirements on their role must be fully supported against the usual tendency to downrate the partisan or critic as "not balanced." Balance is the responsibility of the facts, and not of the good reporter.

We have already said a good deal about the monitor's role, and work load, especially when discussing the Technical Colleague. We would now add three points:

(a) Monitors should certainly go through SP training and as soon as

MP training is packaged, it should also be used in monitor training and--more important still--in monitor selection and promotion.

(Thus we can see pay-off from the push towards systematization, in areas other than improved performance by the primary group. In addition to these gains in secondary group training, selection and promotion, there are other possible gains from remoter clienteles. It may be very useful for Congressional aides, journalists, SEAs and LEAs to go through 'short course' exposure to the training materials.)

(b) The monitors should eventually get their own training program (at least some special sections) for they have special liaison and support duties very different from anyone else in the system.

(c) Their relation to the ST would be very close, since they have much of the most recent data the ST needs for its "War Room"--conversely, the ST has not only the presentations of data but the latest releasable evaluation from MP and SPs, which should be communicated to the monitors.

Short-Term Real-Time Projections. 10/71-6/72:

If an evaluation plan is picked and refined by Christmas, and if it includes major features from this plan, we would see this sequence as optimal:

January	Selection of MP Chairman.
February	Appointment of ST and new Monitors; for 6/72 start at latest.

March	Completion of MP roster; Completion of contracts for training.
April/May	Preparation of materials for MP summer camp. Handling non-postponable decisions. Completion of training program. War Room set-up Instrument contracting or developing.

The Master Panel Approach: An overview: At this stage, enough of the system has been examined to make a general comment appropriate. It is clear that the Santa Barbara team believes:

1. It is too early for objective instruments to be given, or even developed forthwith, that will improve on panels.

2. It is long overdue for us to improve on panels by improving panels. The methodology of doing this is the principal content of this report.

3. It is not too early to start work on instruments that may come to replace, if not panels, at least much of the time needed to calibrate them and convert their discussions into maximally useable form.

4. What counts as an instrument depends on the location of the user with respect to it. A panel is an instrument from the point of view of DR^2 and it can be validated like any other. From its point of view, it is not an instrument and there remains open the question whether it uses one. The suggestion here is that what it does use should develop into one (several) between Year One and Year Two, and into a validated one by Year Three.

5. Whatever system is adopted must be set up/treated as a good experimental design. In particular this means a meta-evaluation system. Because this is almost entirely applicable to any system, it is presented as a separate module, scarcely discussed in the preceding.

Probably the ME should devise some other check-up procedures of which he says nothing in advance, except that there will be some (a condition of membership in the MP and--not necessarily the SPs--would have to be the willingness to have one's own operation examined) and that it will be cleared in advance--but possibly only a few moments in advance--with the chairman of the MP and the head of DR^2 (as well as the presumably secure routine questionaire clearance unit in OE/NIE).

The above-mentioned activities are related to only one of the questions of concern to the ME, viz., validation of the conditions of the MP. And there are many other tests that help to regulate on that, some of which are built into the calibration program, which he will watch carefully and parts of which he may repeat later in the year. (For example, he would be interested in connections between low scores on the calibration tests and deviant positions in later evaluation discussions).

Another important question for the ME is the relative validity of the MP system by comparison with other feasible systems. It may be that the validity is low, but as good as we can get; or high, but readily surpassable. Thus, the ME will try to formulate or discover feasible alternatives for each replaceable component or subsystem of the system, as well as for the whole, and try to get reliabilities on each. For example, he might check on the reliability of the selection of the SPs by the MP + ST, by sending out the basic list of candidates to experts who were not consulted to get their selection; and by giving an operational and replicable description of the way he selected them, he would have a basis for an improved method if his approach gave different results and independent tests show 'his' list to have been better.

The Metaevaluation Subsystem; Without this, or something like it, nothing reliable will ever be known about the merits of the evaluation system used by DR^2. It is cheap to have, very expensive to cut. And it provides an order of magnitude increase in credibility.

The metasystem requires--for Year One at least--one full-time independent professional evaluator, or the equivalent in the services of a team. It should probably be RFP'd in March for contracting in late May, so that the first Summer Camp can be covered. The 10% rule makes good sense here-about $38,000 should cover it, assuming a half-time secretary and a fair amount of traveling.

The metaevaluator ("ME", hereafter) has to cope with both formative and summative roles. He should be encouraged to be open about most of his plans, but not all. Making it clear that he will, for example, conduct telephone quizzes of MP and SP members on proposals they have recently evaluated (to check whether evaluations are based on adequate understanding of what is evaluated), is likely to have salutory effect of the care with which reading is done, by SPs at least (the effect may wear off with the MP). But to say which day this will be done is going too far, of course. To do it at all destroys our base line, but it would be reasonable to suppose that panels as used by other divisions of OE could be taken as representative of how this one would have been without (a) special training (b) the announcement by the ME. To evaluate the worth of the new plan, such a baseline should be obtained and probably another division would be interested in getting this data.

The ME should be able to attend meetings of the MP to study process and occasionally perhaps to make suggestions or to raise questions in his formative role. He would particualrly watch the extent to which MPs work towards making their criteria explicit, since that is one of the charges to them which is not self-checking, unlike "rate these proposals". There might be occasions when he arranges a phony application to go the rounds, possibly as a substitute for the phone quiz, perhaps for other purposes.

The ME would debrief everyone on the MP, a selection of SPs and the ST, at least; on the other end, a selection of unit personnel; and in the middle, the monitors. For formative purposes, he might do a partial job on this in the mid-year.

He should have part of the role of an ombudsman and all participants in the program, including <u>all</u> unit personnel, should be informed of his existence and of the fact that he welcomes anonymous as well as signed comments on the system or suggestions about how to investigate it. Some such comments, he may feel, should be passed on--further coded, perhaps--to DR^2 management. Others may be followed up for accuracy.

It is likely the evaluation will use some secondary indicators and the ME should regard it as a major task to begin the process of validating these. The two most important appear to be the "mission thrust" and the "critical mass" requirement. A good deal of evaluation has been previously done against these requirements, but their validity is quite uncertain. The ME must devise or select ways to test them or avoid them.

One task of importance would be the check on impact with the 'remoter public' of the plan (Hively's interest); to see whether output for the system is reaching them and how credible it seems.

The ME should not have an office in Washington, to reduce co-option pressures...He should run snap checks on conflicts of interest... He should follow-up with the units the utility of MP/SP evaluation... No good evaluation of a new system can be <u>completed</u> in the first year--a follow-up would be important, but could be passed on to the ME's successor. But what a good ME could send in would be of great importance to DR^2 and the chairman of the MP, any factual corrections following later from his successor.

Who evaluates the metaevaluator? Well, for $3800 we could have a graduate student follow him around to see he didn't take any bribes. How about <u>that</u> guy? For $380, we could get a large St. Bernard to follow <u>him</u> around.

V. CRITERIA AND PROCEDURES

This section attends to some more detailed questions about the MP Plan and takes up some general questions that may by now be pressing on the reader. It begins with a statement of an important long-term goal of a good evaluation system, towards which the criteria and procedures of the section are connected.

Symbiotic Evaluation. A good evaluation system for DRDR should, in the long run wither away almost. Good evaluation is so important to the evaluee (e.g., the unit manager) that in the long run he should adopt it as part of his budget. Naturally, he will take extreme precautions to avoid co-opting his evaluators, but there are a number of ways to do this. He wants to satisfy prospective customers for his products that they have merit. He wants to know which of his research teams are having the most impact.

Of course he also wants his funds from OE/NIE to continue and this may tempt him to make it easy on himself. But should this appear from the records, or show up from a spot-check, the prospective loss to himself should be so disastrous as to offset the immediate charm.

Sponsoring agencies will always need--for Congressional credibility at least--to keep one eye on the evaluation procedures of their grantees, which means reading the reports and spot-checking. But the

evaluation should be at least symbiotic, rather than parasitic (or carnivorous as it is too often currently perceived.

The idea of mature institutions doing their own milestone checks (in 6/3/71) is a step in this direction--but a premature step and without safeguards. Premature because the level of sophistication amongst site visitors is far too low now. And the safeguard of spot-checks must surely be retained.

Both formative and summative program evaluation have a role for an institution since it will hopefully outlast any one of its programs; they may both be regarded as part of formative institutional evaluation. The symbiotic role of the DR evaluation system should eventually become that of the institution's own evaluation system.

But to get to this point requires a great deal of education, not so much for the directors of mature labs as for program directors and center personnel--and many evaluators. Where there is some recognition of the point, it tends to be lip-service only, and one must realistically look forward to some years when the spot-checking will be more than nominal.

Nevertheless, one goal of a good evaluation system is to prove its value to the evaluee, and this should be borne in mind as the following procedures and criteria are considered.

The other side of this point is the goal of making OPPE's evaluation of DR^2 redundant--as long as DR^2 is within OE.

Another self-application of interest concerns NIE's projected in-house R&D capability. Here NIE will have a chance to show whether it means what it says about the value of evaluation; and the independent metaevaluation will be a key in determining whether it does. The possibility of using Swedish, British, or German evaluation teams should be explored, since there is a problem for the 'independent' evaluation shops in this country of condemning one of their biggest customers. The problem will be less serious numerically with the creation of NIE, but more serious since NIE's identification with its own research may be expected to be greater than that of OE with 'its' programs.

What kind of data will be needed? This evaluation system is very heavily performance-oriented. Whenever possible, non-pay-off criteria are pointed out. This means some strain on natural tendencies amongst evaluators who notice imperfections in experimental design, fiscal procedures or democratic participation of unit personnel in governance, and wish to fault the unit for them. But it also requires that DR^2 not invoke management mythology in its demands. The Frye document, by and large, shows sympathy with this position and it is known to be Ward Mason's preference. Nevertheless, some deviation from it occurs in DR^2 procedures, some in the AIR product review document, and a great deal in site visitors' reports. The latter must be handled in a calibration program and can be handled extremely quickly by using the good and bad model approach. There is one legitimate exception to performance orientation and that is Phase I-II units when no significant amount of product is available. Then, a critique of management procedures is in order. But it must be recognized as introducing

extra and notably unreliable assumptions and the metaevaluation system should have a provision for checking on these assumptions. The most dubious is the "critical mass" assumption which is quite likely to be a function of the personalities of core staff and the topic, and in no simple way related to absolute numbers.

The most rigid supervision by the chairman of the MP, and by the metaevaluator, will be required to avoid the MP from applying to mature institutions the criteria they must apply to developing ones. But the system proposed does a great deal to guarantee that segregation, as we shall now see.

Forbidden Moves: There are four kinds of question/comment that must not be asked/made in evaluation of the units.

1. "Is the basic program plan of this mature institution, now before us for a milestone check, something that we find philosophically attractive and responsive to our present needs?"

Illicit Because: The BPP has already been reviewed, at the time of funding, and a moral commitment made through completion, barring (a) serious failures (b) extremely serious funding cuts. Even the Rite of Passage evaluation has to swallow this, if it wants the unit because of potentialities and/or other programs in it; and a good case can be made that NIE inherits the obligation to complete what was undertaken in good faith when it is doing well. The cost is not very high, since these 'contracts' only have a few years to run.

Safeguarded By: Milestone review panels are only given a basic match/mismatch task. Does achievement match what was promised for this milestone? If not, is the shortfall/overrun significant? Can

you reliably identify the causes? The MP has no basis for raising further questions if the milestone is passed. If not, and if the examination is not wholly satisfatory, then it legitimately gets into overall review problems, for which it is well prepared by its early training in the total picture. MP & SP training, and especially monitor training, must involve heavy emphasis on match/mismatch work. Its importance for the monitor arises in connection with the present role of aiding unit officers in writing BPPs, which (rightly) require milestones to be stated in a significantly testable way. Knowing how to do the test is a valuable basis for knowing how to write the description.

2. "Is the management of this mature institution performing well by the usual standards of management given that milestones are being passed and the CPA audits (which should be required) are okay?"

Illicit Because: The way mature institutions are managed is the data for books on management science, not a fit subject for evaluation by them.

Safeguarded By: Neither the Ten Decisions nor the elaboration of them herein provides a toe-hold for this question. It would simply be out-of-order.

3. "Since this institution has two programs in bad trouble, we should drop support for the programs and chalk up a debit to unit management; that's what performance-based criteria talk means."

Illicit Because: The reasons for shortfall must be discovered. Given these, the proper action may be a joint DR^2 unit decision to cut the program(s) with no blame to the unit, or it may be full blame for the unit and an override order cancellation of the program, or any mix. There is no inconsistency between saying that success of

programs should not be offset by criticism of management methods against unreliable paradigms, and saying that program failure does not necessarily imply management failure. There is indeed the symmetrical possibility of success despite handicaps imposed by bad management, but the SPs and Technical Colleague network will pick that up in due course--and there should be a lag to allow the many self-correcting mechanisms to operate.

Safeguarded By: Milestone failure or audit/fiscal-procedure-check failure automatically kicks in a SP visit by a management specialist panel. No MP evaluation for program termination or unit discipline can occur until this SP reports.

4. "The Rite of Passage requires a thorough review of each program, hence presumably fifty-odd site visits by SPs."

Illicit Because: Programs that are on target can be judged from their BPP. Institutions can be judged for their track record, if of several years standing. When any doubt exists that they are mature institutions, they can be taken into NIE on Phase II status (possibly Phase I in extreme cases of new and improved institutions e.g., WICHE).

The Mission Capability Criterion: Of all the criteria, this one presents the greatest difficulty. It is resented by the units, ambiguously interpreted by site visitors and not well explained by the agency. The defensible points in it seem to be these:

1. The President or Congress or the Commissioner (Director) or even the MP may identify an educational need for which R&D is required and for which no free or no adequate capacity exits. It seems entirely appropriate that in such a situation they should let a

developmental contract which is keyed to this mission, and hold the contractor to the mission.

2. Ongoing needs or potential needs may require R&D capacity in 'ready' status, even if occupied with current programs (cf, the need for a U. S. chronometrical industry in peacetime.) "Drifting" from this status is a loss which should be corrected or compensated.

3. On the other hand, some of the best-managed units have a nearly random assortment of programs and are not penalized for this, nor is it apparent that they should be, But this creates a sense of injustice in other labs when they are inhibited from expanding in a direction their personnel find attractive.

The best combined strategy that takes account of these points might be as follows:

Phase I and II: mission integrity can be and usually will be part of the contract. But at a certain point in time, there may arise a need for new general purpose centers, now that their possibility has been demonstrated e.g., by CEMREL. The group unit has an important advantage for DR^2, viz., flexibility. Whether it has drawbacks depends on the strength of the hypothesis about "critical mass" which might easily be called questionable.

Phase III: the unit may move towards General Practitioner status as it wishes subject to penalty for any detectable ill-effects on its existing programs.

When DR^2 feels it must have ready capacity in some area, for possible emergency or for some plans it sees coming up, it must pay for it by initiating and providing support for programs in the needed area, or by paying higher fees and etc., for 'narrow' renewals.

46

The Contingency Mix Problem: Suppose the budget for labs and centers is cut 10%. There are various strategies of response.

1. Cut out the lowest ranking programs, from the bottom, until $\geq$ 10% saving is achieved.

2. Do as in one, but skip any that would mean the demise of a center (e.g., because fiscal critical mass is no longer feasible).

3. Prune 10% off all budgets, or the fattest, or 20% of the weakest half, etc.

4. Cut the weakest institution until $\geq$ 10% is achieved.

5. Mixes of the above.

In general, it is doubtful whether any rational basis can be given for preferring one of these to the others. In particular cases, some would be preferable. The MP can help with this, especially by noticing whether any clear preferences exist apart from political pressure. The more open the decision, the better the credibility NIE will retain with respect to the units.

These strategies presuppose the main MP activity, which is grading. Some refinements of this are appropriate.

'Rank and Cut' and other Grading Procedures: The basic grading procedures is semi-ordering against an anchored scale:

A: Outstanding/Excellent

B: Above average/Good

C: Above average/Tolerable

D: Below Average/Weak-Marginal

F: Exceptionally bad/Intolerable

The term before the slash refers to statistical anchoring, the one after to ideal ("absolute") anchoring. Either can be called "normative"; it is usually important to distinguish them in principle, often not so in practice.

This type of grading has a natural "cut" in it at the F/D line, and is very useful for the kinds of evaluation the MP will take on. For the MP, regardless of fund cuts, must always be willing to recommend termination of a program (when the option comes up) on the grounds it is simply not worth funding any more. That it cannot operate on a statistical norm is clear when considering the evaluation of proposals (new units or new labs or new directions). The MP should always be perfectly willing to say that none of these submitted should be funded.

A second type of grading corresponds to the Pass/Not Pass or Satisfactory/Unsatisfatory grades of the academy. Units will be graded in this way with respect to passing an audit, obeying the equal employment requirements, etc., where passing is mandatory; and with respect to availability of adequate local computer facilities, etc., where the absence is a drawback but not fatal, for many programs.

A third type involves "off-scale" or "segregated" or "A+" grading and can be combined with a refinement of the preceding one by adding these conventions:

A+: A breakthrough or performance of such merit or promise that it must be preserved at almost any cost. Sufficient to offset low grades on other scales in the battery, including F (but not F-).

F-: A failure to achieve a satisfactory level on an absolutely necessary requirement.

The great trap in grading is averaging, although it is sometimes the best way to present results. The right way of segregating (a) grades and (b) scales, for a given grading problem is one that cannot easily be reduced to rules at this stage of evaluation methodology (and perhaps never). One MP must have, and the metaevaluator should have, some degree of expertise on this. Grading is implicit in almost all evaluation, and it is the most useful method of representing results--if used properly.

Two comments: (1) When reporting MP or SP decisions to DR^2 or NIE, or the units, a display showing the grade distribution by panelists is the most valuable. For some purposes it helps to give the grade distribution before discussion as well as after. (2) Off-scale grading does not cover all Pass/Not Pass situations. Lack of computer facilities is F- for some programs, not for others. Hence, the scales must be segregated, too; thus "Grade for program of type 1" (F- or NP); "for type 2" (A-D, or P).

Institutional Criteria: All units should be scored on all the following. The way they are graded may involve different weightings for some of these; for example, for some institutions some of the criteria are absolutely necessary; for others, certain criteria are irrelevant, (management criteria for mature institutions).

1.	Research Output	Primary criteria for centers, labs, and developing institutions respectively
2.	Development Output	
3.	Management Performance	
4.	Training Output	Bonus points criteria for all institutions
5.	Dissemination	
6.	National Leadership	
7.	Community Contribution	
8.	Mission Capability	Absolutely necessary for some institutions
9.	Audit	Absolutely necessary for all institutions
10.	Civil Rights Act	

A false impression would be created by exact numerical weightings, but the evaluator should suppose that a 'C' on a primary criterion is absolutely necessary before it's even worth looking at Bonus Points criteria; and that the latter can have a maximum effect of only one grade point.

The instruments for evaluating these criteria can be handled in two ways. They could be RFP'd now or we could operate from the fairly good procedures we now have available in two of these cases, and the obvious, relatively primitive ones in the others, and develop better ones. There is a strong case for the latter approach, since the instruments should be developed for this task and RFPing on the basis of other knowledge seems premature.

Criteria for Research Output: This is of course one of the hard ones. The Santa Barbara team made a serious effort to develop an instrument for this (from a draft of Gene Glass), partly to see whether it was feasible. There was no unaminity about whether it could be done satisfactorily, but it clearly should be tried. While certain conceptual difficulties are apparent (what units for quantity of output;

what comparison base, etc.), it looks as if an instrument might be developed whose imperfections, while serious, may be less than those of the global human response. There has been little recognition of the need for a center to get a formal model of external evaluation of its own research products, but the need is there and the lab-center system should be able to develop such a procedure. The better labs have seen this more clearly. In either case, research <u>is</u> produced - although it is the primary obligation of only one party - and it should be evaluated in the same way for both parties. One would expect that the contribution of a lab to research, or a center to development, would have the status of a bonus points criterion. But there are already cases (possibly Wisconsin) where parity will be attained between the R and the D in R&D. Is such a unit <u>twice</u> as useful? Is one horse that has two gaits twice as useful as two horses each of which has only one? Not if the owner knows what gait he wants on what trip. The crucial criterion is <u>amount</u> of <u>valuable</u> output. And that is not going to be increased by high inter-person variance in type of work, within the unit. (On the critical mass view, it will be hurt.)

This problem with the 'total score' approach is quite general. If visitors notice one lab doing good community relations, they tend to count it against the next that it doesn't, because checkable items are more salient than increments in quality output per person-hour or per $. Yet, one cannot dismiss bonus performance. Hence, the primary emphasis must vary heavily on the main mission (R, or D, or R&D, or Policy), with just the possibility of gain for other considerations.

Research can probably be best judged, in the present state of the art, by an SP which received all documents off-site, which is paid to read them and rate them and send in the rating before convening on-site, where a preliminary discussion would be followed by detailed interaction with individuals or groups. A tentative re-rating session would then occur, and an exchange with the director and team chief. Before leaving, a final rating would occur. It would probably be best if both the submitted ratings were anonymous, though specific comments would be signed. Thus the panelists might avoid the present pressures to be nice which are not conducive to overall reliability. The opportunity to discuss plus the necessity to rate before and after seem to bring out the best in evaluators and also gives the metaevaluator useful data.

More than formal reports must be rated. Centers should be encouraged to submit lists of presentations and consulting activities that may contribute to 5 and 6--dissemination and national leadership. The SP can also comment on the grapevine effects of the work done at the center, for the same criteria.

In moving towards a more formal instrument, it seems clear that comparative ratings with the products of non-NIE centers must be undertaken. Probably the crucial decision for NIE is between money into individual or team basic research grants and money into special centers. NIE needs to know roughly what the cost picture is, between these alternatives, and it can't get them without comparative merit ratings.

The Advocate Team devised several ways to anchor the grading scale, (especially Hively) against paradigms in the field and against 'average research' suitably chosen. It is clear the necessary instrument will require a great deal of work to anchor and more during Year One--but it should be allocated early, perhaps soon after the appointment of the Chairman of the MP. The preliminary suggestions of this Team, while not submitted as part of this report, are in the Project Officer's files.

Criteria for Products: The present situation is both PEP & AIR have produced forms and procedures but no conciliation has occurred. It is easy to snipe at AIR for excessive use of secondary criteria and at PEP for insufficient exposition, but no one-day attack by a third group will produce something better. The best route now might be to put the AIR people to work on a set of products that PEP evaluated, allow them only the same time, have them produce evaluations and test-market these with a sample from the several clienteles of interest. Glass and Schutz on the Advocate Team came up with a sketch-draft of a new attempt, which was discussed at Santa Barbara, and it can also be consulted in the files. One more year's work in this direction will undoubtly yield a truly useful instrument; the prognosis with the evaluation of research probably has to be in terms of 2-3 years. In either case, calibration training is absolutely essential for raters, especially in the area of saleability, which affects cost; and of realistic cost as a basis for merit appraisal. The present median cost level of lab-developed products is surely unrealistic.

Criteria for Management: Here again an excellent basic document is available, developed at great cost by HEW for contractors (see appendix A).

It provides (pp. 69-94) a thorough house-keeping check. More detail is needed in the imaginative dimensions, and specifically educational examples should be developed for training purposes. For a long time we may suppose that the wisdom of experienced R&D administrators, from within and without the DR^2 system, will provide not only useful additional perceptions of merit, but better reliability in applying this instrument. Breaking "staff utilization" into its components, adding criteria for an evaluation system, checking on the extent to which DR^2 is subsidizing a proposal-writing enterprise, or a form-returning one: these are all conclusions for the Management SP and the RFP it should be designing.

Criteria for Training: Intern-hours times quality-multiplier (between 1&2) would be as good as a way to start as any. Of course, the move to performance tests to get the multiplier should be immediate, and the move to drop the hours in favor of a performance gain should be next. BEPD should be leaned on heavily for help (dollars and time and perhaps joint RFPs).

Criteria for Dissemination: Partly covered under research for dissemination of research. Crude measures such as number sold or used would be a good start, together with SP estimates of number of imitators, etc.

Criteria for National Leadership: Something is still left out when all the hardware has been weighed, something which was important in the conception of the educational labs. It is adumbrated in talking of "number of imitators" in the last subsection. We need to pick up some idea of the influence, the leadership, coming from the panels and stimulating new standards and types of work. A preliminary suggestion on this is in an appendix B (Stake).

Criteria for Community Contribution: A holistic grade on this judged as an ancillary activity should be adequate. If, and as, it becomes a more crucial consideration, more complex scales and planning might be called for.

Criteria for Mission Capability: To the extent this can be identified, it can be specified in the initial contract and judges on a match/mismatch basis for those institutions which are bound by it. Unless the validity of refined judgments can be demonstrated, ratings should be restricted to Pass/Not Pass.

Remaining Criteria: Fortunately for the MP, these criteria can be left to someone else.

Membership of the Master Panel: The task is now defined. Who should perform it? The Advocate Team took some early passes at this and then shelved it until the outline of the whole system was on the table, and this report reflects their conviction that selection is facilitated when the task has been fully clarified.

The MP should be picked iteratively. The first move is for senior DR^2/NIE personnel to select a Chairman. With his help, the Support Team and Assistant Chairman. With their help, and the suggestions and constraints below specified, the MP. With their help and that of NCERD's back files, the SPs, especially the Management SP(s). With the aid of Year One's performance data on these evaluators, the key man and drop-offs for the following year. It was felt that the choice of the

one-third carryover for the MP should be made by anonymous ratings of the MP and ST on those available amongst themselves, to increase credibility by comparison with in-house selection.

The following selection-constraints are suggested:

Category	Selection Group	Number	
Elder Statesmen	1 from National Academy of Education	1	
	1 from AERA Council or ex-Council	1	
Report Readers	1 from Boards of Five Best Educational Journals	1	
Educational Technology	1 Educational Technology Expert, top 20 thus selected by direct peer-group ratings or probably rated by peer-groups	1	
Evaluator	As for Educational Technology	2	Year 1
		1	Later Years
Private Sector	Industrial R&D groups: list 10 (IBM, Xerox, BRL, etc.), request their candidates, select for these.	1	
Ex-USOE (non-NIE)	Develop (expanding) list of senior people who are noted for overview	1	
Other agencies	NSF, OEO,BOB etc. Develop list as SPs i.e., identify experience by objective criteria, get peer ratings, selected from top 20.	1	
Sundry	Board of Great Cities Project	1	Year 1
	Historians of Education	2	Later Years
	Critics of Education/Futurists		
	NIE Board		
	Good Deans of Good Education Schools		
	CSSO		

Within these constraints, staff (as defined above) would select, bearing in mind other desirable balances--racial, sexual, etc. The intent is in general not to be representative but to get very good, very knowledgable people who are willing to learn, work and be evaluated for it.

VI. BUDGET:

1. Master Panel

Assuming one unfillable place or one absentee per management or training session, gives ten as cost basis.

Chairman	30,000
Summer Camp 10 x 4,000	40,000
Consultant Fees 10 x 150 x 36	54,000
(second summer, one month; one-time cost, not summed)	20,000
Travel and per diem 9 x 10 x 300	27,000
Office/phone/mail 9 x 200, rounded	2,000
	173,000

2. Support Staff-Central Office Costs

Secretary for Chairman	8,000
Support Team Leader	24,000
2 x (M.A. or B.A. + experience) 13,000	26,000
1 Secretary	7,000
1 Librarian (one-half time) 12,000	6,000
	71,000

3. Specialist Panels/Technical Colleagues/Consultants

50 (programs) x 2500	125,000

4. Auxiliary Costs

Training programs and Instrument Development	50,000
Metaevaluation: staff plus travel and expenses	38,000
	88,000
External Costs: Items 1, 3 and 4 above	386,000

APPENDIX A

REVIEW AND EVALUATION OF MANAGEMENT

REVIEW AND EVALUATION OF MANAGEMENT

A review and evaluation of management--whether formally performed by the Department of Health, Education, and Welfare or undertaken by the organization for purposes of self-improvement or self-assessment--should utilize the following outline of systems, subsystems, objectives, and measures of performance. The measures of performance are illustrative examples of sound practices and procedures which contribute to the attainment of defined objectives.

Although the measures of performance have broad applicability, they are not to be considered obligatory practices since, for any given objective, other procedures (measures of performance) may exist within the organization which just as adequately serve its particular needs.

GOVERNANCE

- The organization's mission is clearly stated and known by all key employees.
- The organizational structure provides for clear lines of authority and responsibility at all levels.
- A written organization chart is prepared and updated when appropriate.
- The organization has a governing board which has the authority to determine or approve major policies and other major actions affecting the organization.
- The organization's financial condition is sound.

SYSTEM, SUBSYSTEMS AND MEASURES OF PERFORMANCE

I. Fiscal Administration (System)
Objective: To provide for control and use of the financial resources of the organization.

A. Budget Administration (Subsystem)
Objective: To provide management with a control mechanism over the utilization of resources in accordance with the approved budget and to assign appropriate responsibility for this control.

Measures of Performance:

1. The grantee organization has an accounting system which identifies cost centers and provides cost data on a timely basis to financial and program managers as a management aid in administration of their programs.

2. There is a means by which management releases the authority to spend under the budget plan.

3. The organization has fiscal controls which result in:

   a. Control of expenditures within the approved program plan.

   b. A management review prior to issuing budget amendments or incurring obligations or expenditures which deviate from the program.

4. There is timely, periodic financial reporting to management which permits:

   a. Comparison of actual expenditures with the budget plan for the same period.

   b. Comparison of revenue estimates with actual revenue for the same period.

5. Responsibility for maintaining budget control is established at all appropriate levels.

6. Analyses and projections are made of cash flow and appropriate action is taken to maintain a favorable cash position.

7. Analyses and projections described in ("6") give appropriate consideration to the use of letters of credit by Federal agencies which minimize cash balances under sponsored programs.

B. Financial Accounting (Subsystem)
Objective: To maintain financial records on a consistent basis in accordance with generally accepted accounting principles for organizations of a similar type.

Measures of Performance:

1. The accounting system, including equipment, meets the needs of the organization.

2. Costs are assembled in a form that meets the needs of the organization.

3. Periodic internal audits are conducted.

4. Corrective actions are taken in response to audit findings.

5. An adequate program for bonding is in use and is reviewed periodically.

II. Procurement (System)
Objective: To obtain the property and services needed in order that the organization may carry out its planned objectives.

Measures of Performance:
A. Written procurement policies consider such matters as cost, quality, delivery, competition, source selection, and subcontract administration.

B. Written procurement procedures cover competitive bidding, negotiation practices, followup on unfilled orders, receiving, inspection, and acceptance.

C. Authority to procure and to sign requisitions is formally established.

III. Property Management (System)
Objective: To control and effectively utilize property required by the organization.

A. Equipment and Supply Management (Subsystem)
Objective: To have equipment and supplies of the desired type, quality, and amount available, without overstocking, and to provide for maintenance.

Measures of Performance:
1. Proposed purchases are reviewed selectively to avoid acquisition of unnecessary or duplicative items.

2. Records are maintained which provide a description of equipment, acquisition cost, and location.

3. A procedure exists to minimize underutilization of equipment.

4. The organization has procedures for purchases of supplies and for central storage and distribution.

B. Salvage, Reassignment, and Disposal (Subsystem)
Objective: To identify and dispose of property no longer required or utilized.

Measures of Performance:
1. The organization has a system for periodic evaluation of equipment and supplies with a view toward economical salvage or disposal of items which are no longer required because of:

a. Obsolescence;

b. Excessive wear;

c. Excessive cost of maintenance;

d. Lack of further need.

2. Authority of effecting disposal or salvage is limited to designated individuals.

IV. Personnel (System)
Objective: To provide personnel to meet program and management needs.

A. Compensation (Subsystem)
Objective: To insure that each employee is properly classified and appropriately compensated in accordance with that classification.

Measures of Performance:
1. The compensation plan includes a scale of rates or ranges based upon responsibilities of each position and its relationship to other positions..

2. Variations in the compensation plan due to differences in requirements, qualifications, and locations are centrally controlled.

3. Compensation rates are not influenced by assignment programs sponsored by outside sources.

4. Fringe benefit plans are reasonable in the context of total compensation when compared to other organizations of similar size and type in the same area.

5. The organization has a policy relating to outside earned income of employees or extent of time spent on nonorganizational activities.

B. Employee Business Expense (Subsystem)
Objective: To provide fair and consistent reimbursement for travel and other necessary business expense.

Measures of Performance:
1. There is a written policy on reimbursement for employee travel expenses.

2. Allowable and unallowable charges are defined for other (nontravel) types of business expense.

C. Nondiscriminatory Practices (Subsystem)
Objective: To provide for equal employment opportunity.

Measures of Performance:
1. Pertinent Federal, State and local orders are followed with respect to equal employment opportunity.

D. Personnel Needs (Subsystem)
Objective: To meet personnel needs through recruitment, orientation and training.

Measures of Performance:
1. The organization has a selection system designed to provide a choice of applicants.

2. Applications for employment (or advancement) in the organization are evaluated on the basis of merit.

3. The organization orients new employees to work assignments, mission, and goals, and to the general standards of conduct expected by the organization.

4. There is a program for staff development.

V. Facilities Management (System)
Objective: To assure that adequate facilities are available for the planned programs and to insure their economical use, maintenance, renovation, and replacement.

A. Space Utilization (Subsystem)
Objective: To assure that space is utilized appropriately and to maximum advantage for the objectives of the organization.

Measures of Performance:
1. The organization has a written policy regarding administration of space.

2. Responsibility is established for assignment and utilization of space.

3. Functions and responsibilities for space allocation are coordinated with those of facilities planning.

B. Health and Safety (Subsystem)
Objective: To assure that the employees and the general public are provided adequate health protection and freedom from hazards.

Measures of Performance:
1. Responsibility is assigned for health and safety matters.

2. There are written policies with respect to health and safety of personnel.

3. A program of insurance coverage exists that provides protection against damage and hazards to the organization, employees, and the public.

4. The organization is in compliance with State and local regulations on health and safety.

5. Where applicable, the organization has a written policy which requires informed consent of human subjects prior to experimentation.

C. Security (Subsystem)
Objective: To provide protection to the organization, its employees, and the general public.

Measures of Performance:
1. There are written policies for protecting employees, the organization, and public visitors.

2. Responsibility is assigned for security.

3. There are controls governing the removal of property.

4. The counsel of local law enforcement authorities is sought to provide and improve the organization's security system.

5. Entrance to restricted areas is controlled.

6. Buildings are adequately secured and periodically checked during off-duty hours.

7. Money handling areas are designed to enhance security.

VI. Planning and Budgeting (System)
Objective: To set forth the nature and scope of each general program (physical, financial, personnel, scientific, etc.) which the organization intends to conduct; to predict the various types and amounts of resources required by the programs; and to arrange for such resources to be available in advance of need.

A. Planning (Subsystem)
Objective: To formulate and adopt a plan based upon priorities for allocation of resources to the various functional entities or programs within the organization.

Measures of Performance:
1. There is organizational planning beyond the immediate year based on stated assumptions.

2. Reviews of budgets and analysis of deviations for previous years show budget plans to be realistic when compared to budget execution for those periods.

B. Budget Formulation (Subsystem)
Objective: To anticipate the financial resources which will be required for an ensuing period.

Measures of Performance:
1. Responsibility for budget preparation is defined at all organizational levels.

2. The final (approved) budget reflects review and consolidation of budget preparations from all organizational elements.

3. The budget process allows time for full review, adjustment, approval, and dissemination prior to commencement of a new budget period.

4. Estimates of revenues and expenditures consider trends established in recently completed budget periods and general economic conditions.

C. Manpower Projections (Subsystem)
Objective: To provide the number and kinds of personnel needed and available to carry out the organization's programs.

Measures of Performance:
1. Responsibility is assigned for planning future manpower needs of the organization.

2. Projections are made of manpower needs and availability for future years.

D. Projection of Facility Requirements (Subsystem)
Objective: To identify the types and amount of building space and fixed equipment available to the various activities to be carried on by the organization; to identify new facilities needed, old facilities in need of renovation, and obsolete ones that should be abandoned.

Measures of Performance:
1. Responsibility is clearly assigned for projecting organizational needs for:

   a. Additional facilities;

   b. Renovation of facilities;

   c. Installation of fixed equipment;

   d. Major equipment;

   e. Disposition of facilities.

2. A capital budget provides for items described in "1."

3. Projections are based on planned program changes.

VII. Management Information (System)
Objective: To provide an information system to meet management needs.

A. Data Collection and Processing (Subsystem)
Objective: To identify the information needs of management and provide a mechanism for obtaining such information.

Measures of Performance:
1. There is an organized method of determining and updating informational needs of management, both program and administrative, both internal and external.

   a. Reporting requirements imposed by management result in fulfillment of these informational needs.

2. Responsibility for projecting informational needs is assigned by management.

3. Projections of informational needs are correlated with long-range planning activities.

4. Management makes periodic reviews of its informational system to insure that it meets current needs.

5. Projections of informational needs are accompanied by recommended means for modifying the current informational system.

B. Reporting (Subsystem)
Objective: To provide timely and complete program, statistical, and financial reporting for internal and external needs.

Measures of Performance:
1. Management reporting requirements are specified and persons responsible for preparing reports are aware of the requirements.

2. The organization provides reports that meet requirements imposed by internal and external sources with respect to:

   a. Completeness and accuracy of data;

   b. Timeliness;

   c. Format.

3. Report requirements are structured so as to avoid duplication of reported information.

C. Records Management (Subsystem)
Objective: To collect, catalog, store, retrieve, distribute, and dispose of management information.

Measures of Performance:
1. Responsibilities are assigned for retention (and disposal) of management information.

2. There are written policies with respect to retention (and disposal) of:

   a. Reports;

   b. Accounting and financial records;

   c. Program data;

   d. Data which become part of the organization's historic record;

   e. Other management information.

3. Filing and storage facilities are adequate, economical, and accessible.

4. The retrieval system is adequate to permit selective data retrieval.

D. Communications (Subsystem)
Objective: To insure adequate transmission of management information at all levels.

Measures of Performance:
1. There is a system for distributing management directives and information to all levels.

2. The system provides for communication upward, downward, and laterally.

3. A series of administrative manuals, accessible to all employees, describes policies and procedures of the organization.

VIII. Inventions and Patents (System)
Objective: To identify and make appropriate disposition of inventions and patents.

Measures of Performance:

1. The organization has a written policy with respect to inventions and patents.

2. Procedures are established which result in notification to management of invention.

3. Formal authority is established to determine disposition of rights.

REVIEW AND EVALUATION OF MANAGEMENT ADDITION GUIDES

The following additional measures of performance will serve as supplementary guides for performing management reviews and evaluations. They illustrate, in greater detail than the previous section, practices and procedures contributing to attainment of system and subsystem objectives.

At this level of detail, procedures for attaining objectives may vary considerably among organizations, depending on their size and type; however, many of the listed procedures have wide applicability.

I. Fiscal Administration (System)

A. Budget Administration (Subsystem)

1. The budget function is carried out separately from that of accounting.

2. The integrity of the budget is maintained by holding amendments and revisions to a minimum during the year, thereby enhancing the budget's effectiveness in cost control.

3. Comparisons of budget with expenditures are made available to department heads and others who have responsibility for controlling costs.

B. Financial Accounting (Subsystem)

1. General

a. Adequate documentation is maintained and is readily accessible to support transactions recorded in the accounting books and records.

b. Unallowable costs (i.e., costs which sponsoring Government agencies do not allow as charges to their grants and contracts) are clearly identified and segregated in the accounting records.

c. Financial statements are prepared at least annually and are presented in conformity with generally accepted accounting principles.

d. The accounting and budget systems are compatible so as to facilitate effective budget administration.

e. The financial statements and financial reports (including reports on individual grants/contracts) are compatible with the budgets covering the same activities so as to facilitate comparison between the statements/reports and the budgets.

f. The financial statements are audited by an independent Certified Public Accountant or independent licensed public accountant.

g. The accounting system is designed to facilitate preparation of financial statements and financial reports.

h. An accounting manual is prepared and is followed consistently.

i. The accounting system provides for consistent identification of direct and indirect costs.

j. Employees in positions of trust are required to take vacations, and their duties are performed by others while on vacation.

k. A double entry accounting system is in use.

l. The books of accounts are kept up-to-date.

2. Processing Invoices

a. Vendors' invoices are compared with copies of purchase orders and receiving and inspection reports.

b. Vendors' invoices for partial shipments are noted on purchase orders to prevent duplicate payment.

c. Invoice computations and account distribution are verified.

d. Transportation bills are audited against purchase orders and material invoices.

e. Invoices for utility services are verified through independent meter readings, records of telephone calls, and other data.

f. Invoices for services, transportation, and utility charges are approved by an authorized official.

g. Debit memoranda are used to charge vendors for shortages, defective materials, etc.

h. Discounts, rebates, debit memoranda, and other allowances are deducted from vendors' invoices before payment.

i. Cost accounts are credited with all charge-backs and allowances made by vendors.

j. Vendors' monthly statements of accounts are reconciled with the accounts payable records.

k. Subcontracts requiring audit are audited prior to final payment.

l. Original invoices are used only for payment or to support payment vouchers.

m. Vouchers are given a final review before signing of checks in payment.

n. Duplicate vouchers and invoices are mutilated to prevent duplicate payment.

o. Checks are mailed without being returned to the accounts payable department.

3. Time and Attendance

   a. Basic work records of the employees are approved by the supervisor or timekeeper.

   b. Overtime work requires supervisory approval.

   c. Attendance of salaried employees is adequately recorded.

4. Payroll Preparation

   a. The payroll is prepared by personnel independent of timekeepers and persons detailed to deliver paychecks or cash to employees.

   b. The accuracy of payroll computations is independently verified.

   c. Payroll totals are cross-checked or reconciled with cost or other department labor summaries for control purposes.

   d. Payrolls are approved by an authorized official of the organization.

   e. All payroll deductions, not required by statute, are evidenced by an authorization signed by the employee.

f. A copy of payroll control sheets or summaries showing total of dollars and hours by departments is sent to cost accounting departments.

g. Payroll preparation is on an exception basis.

5. Payroll Payment

a. The distribution of paychecks or cash is made by personnel who are not involved in timekeeping or payroll preparation.

b. Receipts are obtained when payment is made in cash.

c. Unclaimed paychecks or pay envelopes are delivered to the custody of an authorized official.

d. Payroll corrections and interim and special payrolls are processed in the same manner as the regular payroll.

6. Fixed Assets and Charges for Depreciation

a. Fixed asset acquisitions and retirements are subject to executive approval.

b. Policies for distinguishing between charges to fixed assets and to repair and maintenance accounts are established, clearly defined, and consistently followed.

c. Collateral costs of fixed asset acquisitions are capitalized, including costs of:

(1) Transportation.
(2) Installation.
(3) Initial testing.

d. Additions to fixed assets are recorded.

e. Plant and equipment records are maintained and controlled through general ledger control accounts.

f. Records ("e") are balanced periodically with control accounts.

g. Sales, physical retirements, and abandonments of fixed assets are reported in a routine manner which provides assurance that they will be treated properly in the accounts.

h. Control is maintained over physical assets for which no further use is anticipated to assure the reporting of and accounting for sales or other disposition (including parts and scrap).

i. Fixed assets fully depreciated or fully amortized, but still in use, are carried in fixed-asset accounts.

j. Fully depreciated or fully amortized fixed assets are so identified in the plant, accounting or other records.

k. Depreciation rates are reasonable and computed in accordance with a definite and consistent policy.

l. Depreciation charges are discontinued when an asset or group or assets becomes fully depreciated.

m. A uniform policy is followed in the commencement of depreciation provisions for fixed-asset acquisitions and the cessation of provisions for those disposed of.

7. Petty Cash

a. Responsibility for each fund is placed with one custodian.

b. The imprest system is in use.

c. Petty cash receipt forms are used.

d. Restrictions are placed on types of petty cash disbursements.

e. When the fund is reimbursed, petty cash receipts are marked by the custodian to prevent reuse.

8. Accounts Receivable

a. Accounts are aged periodically for review.

b. Disputed items are handled by someone other than accounts receivable bookkeepers or cashiers.

c. Writeoffs of bad debts and adjustment credits are made only when approved by a responsible official.

d. Credit memoranda are approved by designated personnel.

e. Credit department approval is prerequisite to payment of customer credit balances.

f. Statements are sent monthly to all customers.

g. Delinquent accounts are periodically reviewed by a responsible official.

h. The duties of the accounts-receivable bookkeeper are separate from any cash functions.

i. Where there is more than one accounts-receivable bookkeeper, the account sections for which they are responsible are changed from time to time.

j. Allowances for discounts other than regular terms of sale require specific authorization by a responsible official.

k. The collection department is independent of and constitutes a check on accounts-receivable bookkeepers.

l. Proper control is exercised over bad debts after they have been written off.

9. Cash Receipts

a. Where such items as cash registers, counter sales slips, collectors' receipts, etc., function as proofs of cash receipts, such proofs are checked by an employee independent of the person receiving the cash to determine that the proofs agree with amounts recorded and deposited.

b. Mail is opened by an employee who is independent of persons directly responsible for preparing bank deposits and for posting accounts receivable.

c. A detailed record of receipts is prepared by the employee opening the mail and this record is given to someone other than the employee directly responsible for preparing bank deposits and posting accounts receivable in order to verify amounts recorded and deposited.

d. Receipts are deposited daily, intact.

e. Bank deposits are made by someone other than the person preparing the bank deposit. Neither person has access to customer ledgers or customer statements.

f. A duplicate deposit ticket, after authentication by the bank, is received by an employee independent of the persons preparing and making the bank deposits.

g. Such authenticated deposit tickets ("f") are compared with:

(1) Record of incoming remittances.
(2) Cash book.

h. Negotiable assets, other than currency, checks, or drafts, are held in custody of an employee independent of persons directly responsible for cash receipts and for the maintenance of records relating to negotiable assets.

10. Cash Disbursements

    a. Checks are prenumbered.

    b. Voided checks are kept and filed.

    c. A check protector is used.

    d. A check register is prepared simultaneously with the preparation of the check by mechanical device.

    e. Signing of checks is limited to designated officials whose duties exclude:

        (1) Posting accounting records.
        (2) Recording cash receipts.
        (3) Handling petty cash funds.
        (4) Approving vouchers for payment.
        (5) Payroll preparation.

    f. Supporting data accompany checks when submitted for signature and/or countersignature.

    g. The signing or countersigning of checks in advance is prohibited.

    h. The practice of drawing checks to "cash" is prohibited.

    i. Transfers from one bank to another are under accounting control.

    j. Bank reconciliations are made by persons not involved in disbursement and payroll procedures.

    k. The employee responsible for bank reconciliations obtains the bank statements, or receives the unopened envelopes containing the bank statements, directly from the bank.

II. Procurement (System)

A. General

    1. The organization has written, comprehensive policies regarding the purchase of equipment and supplies.

    2. Responsibility is assigned in a central location for purchasing:

        a. Equipment.
        b. Supplies.

    3. Authority to procure, approve procurement requests, and sign requisitions is specifically assigned.

4. There is advance procurement planning consistent with the size, type, and complexity of the purchasing function.

5. Advantage is taken of discount prices for volume or continuing procurement.

6. Purchasing forms are adequate to support good business practices and special Government requirements.

7. There are procedures for inspection and acceptance/ rejection of supplies and equipment.

8. The procedure for identifying purchase requirements is adequate, and requirements are consolidated when possible.

9. Sufficient procurement lead time is provided by requisitioning units.

10. Reasonable price estimates are given by the requesting unit to the purchasing office whenever possible.

11. Rules are published governing what types of property and services may and may not be purchased.

12. There are periodic inspections for progress and adherence to specifications in construction, alteration and fabrication type procurements.

13. An equipment due-in system is maintained at the receiving point.

14. Subcontract clauses are sufficient to carry out the requirements of the grant/contract.

15. Purchase orders are prenumbered and accounted for by number.

B. Purchasing of Equipment

1. Guidelines are established facilitating decisions with respect to:

a. Manual versus powered equipment.

b. Used versus new equipment.

c. Lease versus purchase of equipment.

d. Utilization of equipment. Use of pooling and sharing techniques.

e. Fabrication in lieu of purchase.

2. Acquisition of major items of equipment is subject to executive approval.

3. Used, obsolete, and other items to be replaced are used as trade-ins when practicable.

4. Purchasing procedures permit compliance with relevant prior approval requirements of awarding Federal agencies.

5. Requests are screened for availability from idle equipment; lists of idle equipment are circulated periodically.

6. Requests for equipment requiring utility services (power, water, etc.) are annotated to assure such facilities are available or will be provided promptly to operate the equipment without delay.

7. Requisitions are promptly canceled when the requirement has been satisfied in another way.

8. Requests for heavy equipment indicate that floor load factors at the place of planned use have been considered and are adequate.

9. Needs are anticipated and requisitions are processed in a timely manner so that items may be procured by normal methods.

10. Requests for equipment are screened to insure that items to be secured are sufficient for doing the job, but not more complex or sophisticated than required to perform the required tasks.

C. Purchasing of Supplies

1. The organization has procedures for determining when to purchase supplies in bulk.

2. Requisitions are processed in a timely manner to assure an orderly flow of supplies.

3. In-house catalogs are maintained to inform using points of items of supply readily available from inventories on hand.

4. There is a central point for receiving requests and ordering supplies.

5. Records are maintained showing types, quantity, and quality of supplies purchased.

6. A supply due-in system is maintained at the receiving points whether bulk purchases or individual item purchases are involved.

D. Use of Competition

1. Adequate competition is obtained on large purchases.

2. When noncompetitive procurements occur, they are supported by adequate written justification from the requisitioning unit.

3. Controls are in effect to minimize noncompetitive procurements.

   a. These controls are placed at a management level sufficiently high so as to be effective.

4. A bidders list or other source list is available in the purchasing office.

5. There is a system of rating subcontractors on performance and such information is used when selecting sources.

6. There is a system of evaluating proposals for technical sufficiency.

7. Price or cost analysis is performed to insure reasonable subcontract prices.

E. Subcontract Administration

1. There is a separate file for each subcontract, containing all documents necessary to support transactions.

2. Procedures are established to monitor subcontract performance, including delivery time.

3. Provisions are made for timely processing of change orders to subcontracts, including price negotiation and revision of the subcontract price.

4. Procedures are in effect regarding inspection and acceptance of items or services received. The purchasing office is kept informed as to completion/progress on the subcontract.

5. Procedures are used to process vouchers for payment to insure timely payment and taking of discounts on subcontracts.

6. Termination clauses are contained in subcontracts to protect the interests of the Government and the grantee/contractor.

7. Adequate procedures are in effect covering subcontract closeout and retirement of subcontract files.

8. There are holdback (or penalty) provisions for subcontractor failure to comply with contract provisions.

9. Procurement documents are processed promptly to the Fiscal and Property Management Departments.

10. There are procedures for adherence to "conflict of interest" principles.

11. There are procedures governing changes in scope, specifications, and cost of subcontracts.

12. There are adequate requirements for insurance and bonding of subcontractors.

13. Procedures are established for handling protests against awards.

14. There is a file of standard clauses to be used when applicable. (Federal, State, and Local Taxes, Use of Convict Labor, Default, Walsh-Healey Act, etc. are examples.)

III. Property Management (System)

A. Equipment and Supply Management (Subsystem)

1. Inventory Control of Equipment and Supplies

a. The inventory system facilitates the achievement of the following:

(1) Identification of kinds and amounts of equipment on hand.
(2) Location of equipment items.
(3) Assignment of responsibility to supervise the inventory and establish frequency and methods of taking the inventory.
(4) Budgeting for replacement of worn or obsolete items of equipment.
(5) Reporting of capital assets.
(6) Screening for equipment on hand.
(7) Reporting of lost or damaged items.

(a) Identification of items lost or damaged (type of equipment or supplies, brand name, serial number, etc.).
(b) Condition (new, used).
(c) Value (cost of repair or replacement).
(d) Investigation to determine circumstances of loss or damage.

(8) Timely action to purge the inventory of items no longer usable or needed for current requirements.

b. There are adequate, secured storage facilities for the types and amounts of items to be stored.

c. Responsibility is assigned for storage facilities.

d. Good housekeeping of storage areas is maintained.

e. Realistic stock levels are established based on consumption and planned need.

f. Stock issues are reviewed periodically to determine trends in consumption.

g. A perpetual inventory control system maintains inventory levels.

h. Controls are provided to prevent obsolescence or deterioration prior to use of short shelf-life items.

i. Physical inventories are made at least annually.

j. There is a stated minimum value for maintaining unit property records.

2. Maintenance of Equipment

a. Adequate procedures exist for periodically inspecting and calibrating or maintaining equipment according to written specifications.

b. Responsibility is assigned for keeping equipment and historical records of:

(1) Periodic routine maintenance.
(2) Major repairs and overhauls.
(3) Emergency repair of equipment.
(4) Equipment that should be replaced due to excessive costs of maintenance.

c. Equipment is secured (keys removed, equipment sealed, or other appropriate action) when not in use.

d. Adequate procedures exist for insuring that only trained personnel use equipment and that equipment and tools are used only for their intended functions.

3. Receiving Equipment and Supplies

a. There is a designated receiving area for incoming property.

b. A due-in file is maintained at the receiving point consisting of copies of outstanding purchase orders.

c. Receiving reports are prepared for property received and amounts are checked against the purchase orders.

d. Incoming property is promptly inspected by qualified personnel to insure:

(1) Material received corresponds with that described in purchase order.
(2) Items are in good condition.
(3) Shortages or substitutes are identified and action taken.

e. When property is rejected, the person requesting the equipment, the purchasing and accounting departments, and other interested parties are notified.

f. There are procedures for returning rejected property.

g. Receiving records are maintained so as to permit ready location of information regarding property received.

h. Items are picked up on inventory as received and as appropriate.

i. Property is distributed to ultimate user or storage room in a timely manner after receiving report is processed.

B. Salvage, Reassignment, and Disposal (Subsystem)

1. Unneeded items turned in by one unit are screened against requests from other units, prior to declaring items as excess.

2. Items determined excess are promptly reported.

3. A determination is made as to whether excess items are suitable for trade-in to reduce cost of items being purchased.

4. Proper authority is obtained for disposition of excess property.

5. Property is disposed of within a reasonable time period after disposal authority is received.

6. Identification is removed from items prior to disposition, if appropriate.

7. When documentation of disposal is complete it reflects authorization, disposal action, date of disposal, and appropriate accounting entries.

IV. Personnel (System)

A. Compensation (Subsystem)

1. Employee Classification

   a. The organization has a position classification system designed to differentiate between levels of responsibility and complexity of work.

   b. Position descriptions are prepared and titles assigned.

   c. Correctness of position description is periodically audited.

2. Compensation Determination

   a. Analyses are made of job requirements.

   b. Total compensation--salaries, fringe benefits and other compensation--is reasonable and comparable to that paid for similar work in the labor markets in which the organization competes for employees.

   c. Periodic reviews are made of the pay scales in an attempt to remain competitive with other organizations.

   d. Rate surveys are conducted periodically for positions requiring similar levels of skill.

   e. Formal policies are established with respect to the following fringe benefits. Benefits are reasonable as compared to other similar types of organizations.

      (1) Retirement
      (2) Health Insurance
      (3) Life Insurance
      (4) Bonus
      (5) Vacations
      (6) Sick leave
      (7) Use of organization facilities
      (8) Disability insurance
      (9) Other (auto insurance, tuition remission, medical care, severance pay, etc.)

3. Performance Ratings

   a. Performance standards are established.

   b. Employees are rated periodically on their performance

   c. Rating results are periodically analyzed and the results used where feasible to consider future pay adjustments.

4. Manpower Control

   a. Policies are established with respect to working hours and absences.

b. Work schedules, vacation schedules, and other absences are administered consistently.

c. Attendance records are maintained for all employees.

B. Employee Business Expense (Subsystem)

1. Authority to approve travel requests is established at all organizational levels.

2. Travel vouchers are submitted to support expenditures claimed.

3. Reimbursement claims are processed promptly.

4. Items of expenditure are reviewed for allowability and reasonableness.

C. Nondiscriminatory Practices (Subsystem)

1. The organization has a method of evaluating the effectiveness of the nondiscriminatory compliance or enforcement program.

D. Personnel Needs (Subsystem)

1. Recruitment

a. Position requirements are analyzed.

b. Position specifications are developed.

c. Sources of potential employees are identified, giving attention to present staff, and advertising to known sources.

d. Position controls are established consistent with the budget plan.

2. Selection

a. Applicants are interviewed and tested.

b. References are investigated.

c. An evaluation of the applicant is made, using an established, rational procedure.

d. Final selection is made at similar levels in all organizational units.

3. Orientation

a. New employees are given an orientation.

b. Training requirements are established for new employees.

c. A followup is made of new employee performance.

4. Promotion and Transfer

a. Position requirements are compared with employee qualifications.

b. Testing of employees is performed when appropriate.

c. An evaluation is made prior to promotion.

d. Training requirements are determined and training provided for employees in new positions.

e. There is a followup of employee performance in new positions.

5. Separation

a. Exit interviews are conducted.

b. Periodic analyses are made of employee turnover.

6. Organizational Structure and Personnel Requirements

a. Key position requirements are established and periodically reviewed.

b. The organizational structure is formally established and periodically reviewed.

c. Organizational changes are recommended when appropriate.

d. Personnel requirements are anticipated.

7. Personnel Development

a. Employee performance standards are established.

b. Individual programs are established for personnel, including:

(1) Secretarial, clerical and administrative skills;
(2) Professional development through seminars and formal programs;
(3) Career counselling.

8. Training

a. The organization has a training capability to provide training programs for upgrading skills of personnel.

b. Costs of training and development of employees are borne largely by the organization.

c. Training programs are available for all classifications of employees relating to job performance and personal development.

d. There is a program for supervisory training.

e. An evaluation of training programs and training results is made periodically.

9. Use of Consultants

The organization has a policy on the use of consultants which requires that before services are obtained evidence is provided that:

a. The services of a consultant are required.

b. A selection process has been employed to secure the most qualified individual available, considering the nature and extent of services to be provided.

c. The fee is reasonable, considering the qualifications of the consultant, his normal charges, and the nature of the services to be provided.

E. General

1. Personnel Relations

a. Communications

(1) Channels are established for distributing information to employees on a regular basis.
(2) Grievance procedures are established and known by employees.
(3) The organization encourages employee submission of ideas for improving the organization and provides recognition to the employee submitting an acceptable idea.

b. Employee Discipline

(1) Standards governing employee conduct and conflict of interest are established and communicated to employees.
(2) Disciplinary measures are established for conduct violations.

2. Employee Services

a. Medical Services

(1) New employees are given physical examinations.
(2) Educational information on health matters is made available to employees.
(3) Emergency medical treatment is available to employees.

b. Personal Services

(1) Cafeteria services and vending machines are available to employees.
(2) Informational assistance is provided to new employees with respect to housing and transportation.

c. Educational Services

(1) Off-duty-hours courses are provided.
(2) Scholarship or loan support is provided for employees and their children.

d. Legal and tax services are available to employees.

V. Facilities Management (System)

A. Space Utilization (Subsystem)

1. There is an inventory record of building and room space that is periodically updated.

2. The inventory record is developed to show categories and subcategories that clearly distinguish types of space, for example, classroom, library, residence, office, general purpose, medical care, and library.

3. The inventory record describes space in terms of:

a. Building Characteristics

(1) Condition, e.g., of plumbing, roofing, general interior, air conditioning;
(2) Cost (including portion borne by other parties), current value, insurable value, or replacement value;
(3) Functional design.

b. Room Characteristics

(1) Function;
(2) Organizational unit;
(3) Condition;
(4) Square feet of area.

4. Reasonable space standards are established for optional utilization. They are based on:

a. Square feet of space per person, or

b. Square feet of space per activity, or

c. Hours of space utilization per year, or

d. Some combination of the above, or

e. Other rational methods.

5. The inventory record describes the function and utilization of each space unit and provides a basis for comparison of units and a basis for reassigning space within and among organizational units.

6. The inventory record serves to assist in the preparation of:

a. Operating budgets.

b. Capital budgets.

c. Long-range planning and budgeting.

d. Depreciation schedules.

e. Maintenance requirements.

f. Housekeeping requirements.

7. Space administration provides for periodic "down time" required for painting, repair, and remodeling.

8. The inventory record provides information on unit elements of space cost which is used in setting the rates for services.

9. Coordination is established among facilities management personnel responsible for long-range budget and planning activities.

10. The authority for space administration is carefully established so as to insure that space allocation is carried out impartially and in accordance with management and program goals.

B. Health and Safety (Subsystem)

There are policies with respect to health and safety of personnel including such items as:

1. Food handling, preparation, serving, storage;

2. General materials handling;

3. Handling of dangerous materials;

4. Construction and building safety;

5. Use of tools, machines, vehicles;

6. Fire and evacuation;

7. Immunization as appropriate.

C. Security (Subsystem)

None.

VI. Planning and Budgeting (System)

A. Planning (Subsystem)

1. The chief executive of the organization has a key role in planning.

2. A network for planning has been established and is known to all personnel having management responsibility.

3. Each person having management responsibility has a decisionmaking role in the planning process.

4. Plans cover the minimum program development period of the organization (3 to 25 years).

5. Current plans have been adopted for the minimum program development period.

6. Each person having management responsibility knows the key elements of the plans pertinent to his operations.

7. Plans are used to guide operations.

8. Objectives are stated in terms of their effectiveness on the organization's basic mission.

9. Operations are scheduled, using critical tasks as checkpoints toward achievement of objectives.

10. Management information requirements are set forth in terms of:

a. Description.

b. Sources and frequency of input.

c. Format and schedule of routine reports.

d. Immediate retrieval and display requirements.

e. Actions to be taken by specified program persons

on the basis of data which is specified but is not a part of any routine report.

f. Linkage to other data systems.

11. Quantity of performance by production individuals is monitored for each payroll period, or less, by the management information system.

12. Quality of performance of production individuals is monitored by the management information system.

13. Progress toward the achievement of checkpoints and the achievement of objectives is monitored by the management information system.

14. Plans are reviewed and updated at least annually.

15. Plans are reconsidered in detail every 3 years or less.

B. Budget Formulation (Subsystem)

1. All persons having management responsibility

a. Have a decisionmaking roll in budget formulation.

b. Know the key elements of all current and projected budgets which relate to their operation.

c. Consider all current and projected budgets which relate to their operation to be reasonable for the achievement of stated checkpoints and objectives.

C. Manpower Projections (Subsystem)

1. All persons having management responsibility

a. Have a decisionmaking role in manpower programing.

b. Know the key elements of all current and projected manpower programs which relate to their operations.

c. Consider the current and projected manpower programs which relate to their operations to be reasonable for the achievement of their objectives.

2. Manpower requirements are projected by category based on the objectives cited in the plans and the performance standards cited in the budget formulation process.

3. Manpower utilization analyses are performed periodically for each category of manpower.

4. Supply-demand analyses have been made or acquired annually or more frequently for each category of manpower with which the organization is concerned.

D. Projection of Facility Requirements (Subsystem)

1. All persons having management responsibility

a. Have a decisionmaking role in facilities programing.

b. Know the key elements of all facilities programs which relate to their operations.

c. Consider the current and projected facilities programs which relate to their operations to be reasonable for the achievement of their objectives.

2. Facilities requirements are projected by type, based on the objectives cited in plans and stated relationships between facilities and performance/production as established in the budget formulation process.

VII. Management Information (System)

A. Data Collection and Processing (Subsystem)

1. Information in the management information system is compiled in basic or prime units, thereby providing a basis for manipulating data to derive combinations of complex data to meet analytical and operating needs of the organization.

2. The sum of the basic information units constitutes a reasonably complete picture of the organization, i.e., its resources and activities.

3. All units of data within the organization form a comprehensive bank of information for common use.

4. Definitions and categories of information are standardized within different departments of the organization and are compatible within functional groupings.

5. The information system provides status information on the following items to facilitate management control and decisionmaking:

a. Total resources of the organization:

(1) Income
(2) Personnel
(3) Facilities
(4) Equipment

b. Resources allocated to particular segments of the organization.

c. Allocation of resources to distinct cost elements for performance measurement purposes.

d. Space utilization.

e. Organizational output:

(1) Publications
(2) Reports
(3) Services
(4) Products

f. Major organizational activities.

6. Information can be retrieved to show data by the following classes:

a. Historical

b. Current

c. Projected

7. Information can be retrieved to show direction and rates of movement within a time period.

8. If the information system is computer based, it includes the following:

a. A computer development plan, which includes:

(1) Extension of services to all basic clerical functions within the organization;
(2) Integration of as many operating systems as possible;
(3) Common data formats and codes for the same information wherever used, e.g., name and address, employee identification, in computer files;
(4) Systems and programing standards covering such items as language and programing techniques, common subroutines, standard file structures, and program design points for hardware device utilization;
(5) A system for allocation of full costs of computer support to using administrative functions or areas;
(6) Procedures for determining priorities for systems development work, and for measuring the cost effectiveness of new computer systems as they are implemented.

b. Management level computer systems support plans which include:

(1) Ability to merge data from different functional files, such as personnel and financial, in order to produce valid and comparable reports based on unit costs or other resource allocation methods;
(2) Use of remote terminals and online files to provide immediate access to high priority files;
(3) Integration of computer files into a common physical format, either on magnetic tape or disc, to facilitate shared use of common information among different functional areas;
(4) Implementation, to the extent practicable, of advanced management techniques such as exception or "action" reports, which are automatically produced as a result of data values exceeding present limits.

9. Information is updated so that new and current data may be quickly retrieved whenever necessary.

10. The data bank contains readily accessible, up-to-date information.

11. The following rules are normally followed to make the most of the data bank or collection:

a. All characteristics associated with each element of data are stored along with their names (including the files where it is stored).

b. The characteristics of every element of data are perpetually updated and available to the total computerized system.

c. Each element of data can be entered into the system by one and only one transaction type.

d. Every transaction type and file-record type is described in terms of the associated names for the elements of data that make it up.

e. All records and transactions are treated the same way; each type is given its own identification code.

f. Perpetual descriptions are maintained for every record type and transaction type and are available to the system.

g. Every structural change to a file is reflected as a parametric input, and there is never a need to restructure data.

h. The files are data independent and the need to reprogram due to structural changes is totally eliminated.

B. Reporting (Subsystem)

None.

C. Records Management (Subsystem)

None.

D. Communications (Subsystem)

1. Decisions and policy changes due to the present Management Information System are being communicated and implemented.

E. General

1. Management makes a formal determination based on comparable cost benefit relationships as to types of information to be included in the system.

2. The Management Information System facilitates management decisions.

3. The present Management Information System is capable of adjusting to a changing environment within the organization.

VIII. Patents and Inventions (System)

A. The organization has a formal invention and patent policy with definitions of important terms.

B. The policy clearly indicates which employees and students are covered and the circumstances requiring reports of inventions.

1. The organization requires the signing of an agreement form under which persons covered by the policy agree to abide by its terms, including the signing of all required assignments and licenses.

2. The organization has a report of invention form to facilitate the reporting of intentions and acquire all information necessary before filing a patent application.

3. The policy statement identifies the control point to which invention reports are to be submitted.

C. The policy statement provides guidelines for disposition by an organizational body of reported inventions in any one or combination of the following circumstances:

1. When the invention bears a direct relation to or is made in connection with the inventor's official duties.

2. When the invention is made during working hours.

3. When the invention is made with a contribution of organization facilities, equipment, materials, funds, or information, or of time and services of either organization employees or faculty on official duty.

4. When the invention is made in performance of either private or Government sponsored research.

5. When a combination of the above circumstances is applicable.

D. In situations where disposition requires assignment of the invention to the organization, the organization employs legal and technical service personnel either directly or through consultant services to:

1. Perform patent searches.

2. Prepare patent applications and necessary accompanying documents.

3. Negotiate and prepare license agreements.

E. The statement of patent policy permits payment of a percentage of royalty income to the inventor as an incentive to cooperating in reporting inventions promptly, signing all documents necessary during prosecution of patent applications and aiding licensees with technical problems.

APPENDIX B

A PLAN FOR EVALUATING THE EXTRAMURAL STIMULATION AND IMPROVEMENT OF RESEARCH AND DEVELOPMENT BY OE-SUPPORTED LABS AND CENTERS

Robert E. Stake
8/19/71

A plan for evaluating the extramural stimulation and improvement of research and development by OE-supported Labs and Centers

It is seen to be the responsibility of all Labs and Centers to contribute to the larger research and development community within education, not just by using exemplary inquiry procedures and by distributing high quality products, but also by facilitating the conduct of research and development elsewhere by

a. identifying information gaps and product needs
b. encouraging new approaches and innovative combinations of talent
c. training research and development personnel
d. contributing to the informal networks of idea exchange
e. promoting the program and affairs of professional associations
f. providing consultative services to educational and administrative groups as well as R&D groups

It is not supposed that an institution or program can be justified mainly on these activities; nevertheless they should not be overlooked in an assessment of its overall worth.

The measurement of this attribute of R&D institutions--for the purpose of OE quality control--should depend not only on rumor and reputation, but should depend on routine and standardized procedure. The agency

responsible for evaluating the labs and centers should designate a person to be responsible for coordinating and upgrading the quality of this information. It is recognized that research facilitation is difficult to measure directly but it is not unreasonable to suppose that actual facilitation is highly correlated with perceptions of facilitation. These perceptions can be collected objectively and routinely.

The "facilitation" propensity of the institution would be determined partly by looking at the institution itself, partly at a sample of its members. These members would be compared to university and state department personnel. Perhaps half of the comparison group would be persons holding at least a masters degree who have published twice within the past two years (publications of a research, development, administrative or professional nature all qualify). Ratings of personnel would not be obtained for individual members but would be summarized for institutions.

The institution characteristics would be judged in an absolute sense, with attention diverted at times to comparison between OE labs and centers.

The characteristics might be measured in the following way: Information gaps, Product Needs. Proposals and reviews of the literature, product reviews are judged as to their perceived "facilitativeness".

New Approaches. The style of behavior of a sample of lab persons is examined, through interview of those who work with them; also anecdotal

records are collected.

Training. Records are kept of internships, time and funds allocated for training new (particularly young) staff members. Note is made of those leaving lab/center as to whether or not they are better prepared to deal with (next assignments) in research and development than when they came.

Networks. Interviews to find out if staff members exchange preprints, memos, clippings with co-workers outside the lab/center.

Professional Associations. Contributions to programs are tallied. Officers are asked to rate persons on quality of participation.

Consultation. Tapes made of consultation are scored for the quality of advice, colleagues are interviewed.

The following matrix might be followed. Draw 6 persons from each.

	Gaps Needs	New Appro.	Train-ing	Net-work	Prof. Assoc.	Consul-tation
R&D Center						
1						
2						
3						
4						
Lab						
1						
2						
3						
4						
State Dept.						
1						
2						
3						
4						
University						
1						
2						
3						
4						